Wavelets and Singular Integral Operators

by

LAU Shui-kong, Francis

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Master of Philosophy
in
Mathematics

July 1999

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my supervisor Prof. Lau Ka Sing for his aid, comments, encouragement and stimulation throughout the whole period of my postgraduate study and on the preparation of this thesis. I also want to thank Prof. Leung Chi Wai for his advice of my postgraduate study. Finally, I would like to hearty thank everyone who has concerned on my thesis.

Abstract

Abstract of thesis entitled:

Wavelets and Singular Integral Operators

Submitted by LAU Shui-kong, Francis

for the degree of Master of Philosophy

at The Chinese University of Hong Kong in July 1999.

Wavelets were introduced in the beginning of the 1980s. They have attracted considerable interest from the mathematical community. This subject was influenced by ideas from pure and applied mathematics (harmonic analysis, functional analysis, approximation theory, numerical analysis, fractal set etc.), engineering (image processing, subband coding) and physics (coherent states, renormalization groups).

Wavelets provide us with orthonormal bases for $L^2(\mathbb{R})$ that are particularly natural when dealing with the analysis that involves the action of translation, dilation and Fourier transformation. A function $\psi \in L^2(\mathbb{R})$ is an orthonormal wavelet provided the system $\{\psi_{j,k} : j, k \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(\mathbb{R})$, where

$$\psi_{j,k}(x) := 2^{-j/2}\psi(2^{-j}x - k) \ , \quad \forall j, k \in \mathbb{Z}.$$

This system is generated from one function ψ, by translating it and dilating it. The factor $2^{-j/2}$ is a normalization constant.

When one looks at the above definition of wavelet for the first time, one question arises immediately: do wavelets exist at all? In Chapter 1, we discuss a general method that were introduced by Mallat and Meyer for constructing wavelets, namely the multiresolution analysis (MRA). We apply this method to obtain Daubechies wavelets as examples of compactly supported wavelets. We

also consider the MRA in $\mathbb{R}^n$, the situation is more complicated, if one makes certain natural assumptions, it can be shown that one needs $2^n - 1$ of the generating functions ψ to form such basis. An extension of the orthonormal basis is the concept of frame in $L^2(\mathbb{R})$. Furthermore, properties of frame will also be included in this chapter for later use.

Numerical algorithms designed for application of integral operators to vectors will be introduced in Chapter 2. As is well known, a direct multiplication of a dense $N \times N$ matrix to a vector required roughly N^2 operations, and this simple fact is a cause of serious difficulties encountered in large-scale computations. Most iterative methods for the solution of systems of linear equations involve the application of the matrix of the system to a sequence of recursively generated vectors, which tends to prohibitively expensive for large scale problems. We will describe a method [1] for the fast numerical application to arbitrary vectors of a wide variety of operators. The method normally requires $O(N)$ operations, and is directly applicable to all Calderón-Zygmund and pseudodifferential operators. In general, the scheme in this chapter can be viewed as a method for conversion of dense matrices to a sparse form.

Actually, chapter 2 provides two schemes for the numerical evaluation of integral operators. The first one is a straight forward realization ("standard form") of the matrix of the operator in wavelet basis. This scheme is of order procedure $N \log N$. However, the second realization ("non-standard form") leads to an order N scheme, which will be described in this chapter in more details. It is a folklore that Calderón-Zygmund operators have almost diagonal matrices in orthonormal wavelet bases. However, Yang [41] showed that this statement is not true as stated. In contrast, the non-standard matrix representation of Calderón-Zygmund operators always yields almost diagonal matrices.

In Chapter 3, we will discuss the $T(1)$-Theorem of David and Journé [9]. The original proof of this theorem require the well-known near-orthogonality lemma of Cotlar and Stein. However, one can apply wavelet theory to simplify the proof of the $T(1)$-Theorem. Moreover, we will state some recent results on the $T(1)$-theorem with weaker kernel conditions.

In the final chapter, we will discuss the two results of Heil, Ramanathan and Topiwala [17]. Both results are arising from the Weyl correspondence. The Weyl correspondence is a formalism that bijectively associates to any continuous linear operator $L : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ a distributional symbol $\sigma \in \mathcal{S}'(\mathbb{R}^n)$. The first result investigates the asymptotic decay of the singular values of compact operators arising from the Weyl correspondence, the target is to obtain sufficient conditions on a symbol which ensure that the correspondence operator has singular values with a prescribed rate of decay. The second result is an improvement of the Calderón-Vaillancourt Theorem in the context of the Weyl correspondence. The usual Calderón-Vaillancourt Theorem states that the pseudodifferential operator L_σ is a bounded operator on $L^2(\mathbb{R}^n)$ if $\sigma \in C^{2n+1}(\mathbb{R}^{2n})$. The improved version states in terms of the Hölder-Zygmund classes $\Lambda^s(\mathbb{R}^n)$, which says L_σ is a bounded linear operator on $L^2(\mathbb{R}^n)$ if $\sigma \in \Lambda^s(\mathbb{R}^{2n})$ with $s > 2n$. The idea of the proof is to expand the symbol σ in terms of a Gabor frame generated by a Gaussian function and use the fact that the partial sums of this expansion are naturally associated to finite rank operators.

摘要

小波的成形與發展是 80 年代才開始的，它的興起吸引了數學界的關注，而小波分析的構思亦是來自多個方向，如純數學和應用數學（調和分析、泛函分析、逼近理論、數值分析和分形學等）、工程學（subband coding 和圖像處理）及物理學（相干狀態和重正規化群）等。

在談到小波的時候，人們的第一個話題常常是“什麼是小波？”。小波是指一函數 ψ 經過伸縮與平移後所產生的 $L^2(R)$的一個基底。若我們定義 $\psi_{j,k}(x)=2^{-j/2}\psi(2^{-j}x-k)$ 的話，則

$$\{\psi_{j,k}:\ j,k\in Z\}$$

就是 $L^2(R)$的正交基。

當人們第一次看見小波的定義時，很自然會提出一個問題，“小波是否存在？”。於第一章，我們將會討論 Mallat 及 Meyer 在 R 上對小波的構作方法，名叫多分辨分析(MRA)，我們亦會討論在 R^n 上的 MRA，這比 R 上的 MRA 複雜。在 R^n 上，我們需要 2^n-1 個函數 ψ 才能生成 $L^2(R^n)$上的正交基。

在第二章中，我們將會討論一些計算奇異積分的數值算法。眾所周知，若一 NxN 稠密矩陣乘上一向量，大概需要 N^2 個運算，這對大型的計算構成重大的障礙及困難。而大多解線性方程組的方法都用上疊代法，但疊代法一經使用，人們便會碰到以上的問題。在這章中，我們會討論一數值法 [1] ，利用小波去簡化計算過程，而這方法通常只需要 O(N) 個運算，更可以直接運用到 Calderón-Zygmund 算子及偽

微分算子之上。我們可把這方法看成為將稠密矩陣轉成為稀疏矩陣之方法。

在第三章中，我們會討論 David 及 Journé [9] 的 T(1) 定理。David 及 Journé 的證明運用到 Caltar 及 Stein 的著名的近正交引理，但若我們運用小波去證明此定理，證明的過程會簡化得多。另外，我們還會列出數個有關 T(1) 定理的新結果，它們均減弱了對核的條件。

在第四章中，我們討論 Heil, Ramanathan 及 Topiwala [17] 的結果，它們都是由 Weyl 對應所引發的。第一個結果研究由 Weyl 對應引發出的緊算子的奇異值的漸近衰變，目的在於獲得充足的條件確保σ生成的算子L_σ的奇異值有一定的衰變率。第二個結果是在 Weyl 對應意義下的 Calderón-Vaillancourt 定理。典型的 Calderón-Vaillancourt 定理指出若$\sigma \in C^{2n+1}(R^{2n})$，則偽微分算子 L_σ是有界的，但改進後的 Calderón-Vaillancourt 定理則運用 Hölder-Zygmund 類 $\Lambda^s(R^n)$ 來說明L_σ的有界性如下：若$\sigma \in \Lambda^s(R^{2n})$ 及 $S>2n$，則L_σ是$L^2(R^n)$上的有界線性算子。證明的想法是把σ展開成由 Gaussian 函數所生成的 Gabor 標架及用上其部份和對應一有限秩的算子。

Contents

Chapter 1

General Theory of Wavelets

1.1 Introduction

For any complex-valued function ψ on $\mathbb{R}$ and for any $j, k \in \mathbb{Z}$, we let

$$\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k) \quad . \tag{1.1}$$

It is clear that $\psi_{j,k}$ is the dilation by 2^{-j} and translation by $k2^j$ of $2^{-j/2}\psi$. A function ψ is called an orthogonal wavelet if $\{\psi_{j,k} : j, k \in \mathbb{Z}\}$ is an orthonormal basis of $L^2(\mathbb{R})$. For such basis a function $f \in L^2(\mathbb{R})$ can be represented in terms of $\{\psi_{j,k}\}$ as

$$f(x) = \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} c_{j,k}\psi_{j,k}(x) \quad , \tag{1.2}$$

where the coefficients $c_{j,k}$ are given by

$$c_{j,k} = \int_{\mathbb{R}} f(x)\overline{\psi_{j,k}(x)}dx \quad . \tag{1.3}$$

Recently there are great amount of work to analyze functions $f \in L^2(\mathbb{R})$ by using the decomposition (1.2). The coefficients $c_{j,k}$ can be made computationally efficient by additional assumptions on ψ.

In this chapter, we will introduce the notion of Multiresolution Analysis and discuss how to construct orthogonal wavelets, in particular, the compactly supported wavelets. They will be used in the following chapters.

Historically, the orthonormal bases of wavelets were first constructed by Stromberg [38] and then by Meyer [23]. Later, the notion of Multiresolution Analysis was introduced by Meyer [24] and Mallat [22]. Orthonormal bases of compactly support wavelets with vanishing moments were constructed by I. Daubechies [7].

1.2 Multiresolution Analysis and Wavelets

There are many possible choices of ψ which make $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ an orthonormal basis of $L^2(\mathbb{R})$. They range from the *Haar* Wavelet, which is a step function on $[0,1]$, to infinitely differentiable function with support on $\mathbb{R}$. Intermediate possible choice also exists, where ψ is m times differentiable with compact support. All interesting wavelets bases are associated with a Multiresolution Analysis.

Definition 1.2.1 *A Multiresolution Analysis (MRA) consists of a sequence $\{V_j\}_{j\in\mathbb{Z}}$ of closed subspaces of $L^2(\mathbb{R})$ satisfying*

(i)

$$\cdots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots \tag{1.4}$$

(ii)

$$\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R}) \tag{1.5}$$

(iii)

$$\bigcap_{j\in\mathbb{Z}} V_j = \{0\} \tag{1.6}$$

(iv)

$$f(\cdot) \in V_{j+1} \ \textit{ if and only if } \ f(2\cdot) \in V_j \quad , \quad \forall j \in \mathbb{Z} \tag{1.7}$$

(v)

$$f(\cdot) \in V_0 \ \textit{ if and only if } \ f(\cdot - k) \in V_0 \quad , \quad \forall k \in \mathbb{Z} \tag{1.8}$$

(vi) There exists a function $\phi \in V_0$ such that

$$\{\phi(\cdot - k) : k \in \mathbb{Z}\} \tag{1.9}$$

is an orthonormal basis for V_0.

(1.7) and (1.9) together imply that $\{\phi_{j,n} : n \in \mathbb{Z}\}$ is an orthonormal basis for V_j for all $j \in \mathbb{Z}$. Since $\phi \in V_0 \subset V_{-1}$ and the $\{\phi_{-1,n} : n \in \mathbb{Z}\}$ are an orthonormal basis in V_{-1}, we have

$$\phi(x) = \sqrt{2} \sum_n h_n \phi(2x - n) \quad , \tag{1.10}$$

and the Fourier transform of ϕ satisfies

$$\widehat{\phi}(\xi) = \frac{1}{\sqrt{2}} \sum_n h_n e^{-in\xi/2} \widehat{\phi}(\xi/2) \quad . \tag{1.11}$$

The convergence in either sum is in the L^2-sense. The equation in (1.10) is called a dilation equation and ϕ is called a scaling function. Note that

$$h_n = \langle \phi, \phi_{-1,n} \rangle \text{ and } \sum_{n \in \mathbb{Z}} |h_n|^2 = 1 \quad . \tag{1.12}$$

We write (1.11) as

$$\widehat{\phi}(\xi) = m_0(\xi/2) \widehat{\phi}(\xi/2) \quad , \tag{1.13}$$

where

$$m_0(\xi) = \frac{1}{\sqrt{2}} \sum_n h_n e^{-in\xi} \tag{1.14}$$

is a trigonometric series.

Let P_j denote the orthogonal projection operator onto V_j, then (1.5) ensures that $\lim_{j \to -\infty} P_j f = f$ for all $f \in L^2(\mathbb{R})$.

For every $j \in \mathbb{Z}$, we define W_j to be the orthogonal complement of V_j in V_{j-1}, then we have

$$V_{j-1} = V_j \oplus W_j \text{ and } W_j \perp W_{j'} \text{ if } j \neq j' \quad . \tag{1.15}$$

It follows that, for $j < l$,

$$V_j = V_l \oplus \bigoplus_{k=0}^{l-j-1} W_{l-k} \quad . \tag{1.16}$$

By (1.5) and (1.6), we see that

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j \quad , \tag{1.17}$$

which is a decomposition of $L^2(\mathbb{R})$ into mutually orthogonal subspaces, with the property that

$$f \in W_j \ \text{ if and only if } \ f(2^j\cdot) \in W_0 \quad . \tag{1.18}$$

The following theorem in [8] tells us how to construct an orthogonal wavelet ψ if a MRA is given.

Theorem 1.2.2 *Let $\{V_j\}_{j\in\mathbb{Z}}$ in $L^2(\mathbb{R})$ be a MRA, then there exists an associated orthonormal wavelet basis $\{\psi_{j,k} : j,k \in \mathbb{Z}\}$ for $L^2(\mathbb{R})$ which generates W_j such that*

$$\widehat{\psi}(\xi) = e^{i\xi/2}\overline{m_0(\xi/2+\pi)}\widehat{\phi}(\xi/2)$$

(with m_0 defined by (1.14)), or equivalently

$$\psi(x) = \sqrt{2}\sum_n (-1)^{n-1}\overline{h_{-n-1}}\phi(2x-n) \quad , \tag{1.19}$$

(with the convergence in L^2-sense) and

$$P_{j-1} = P_j + \sum_k \langle \cdot, \psi_{j,k}\rangle \psi_{j,k} \quad . \tag{1.20}$$

The most elementary example of MRA is obtained from a *Haar system.* The scaling function in this case is the characteristic function of the interval $[0,1)$ and is denoted by $\chi_{[0,1)}(x)$. Let $\phi(x) = \chi_{[0,1)}(x)$, it satisfies the dilation equation

$$\phi(x) = \phi(2x) + \phi(2x-1) \quad ,$$

the closed subspaces V_j of $L^2(\mathbb{R})$ satisfy (1.4) to (1.9) is given by

$$V_j = \left\{ f \in L^2(\mathbb{R}) : f|_{[2^j k, 2^j(k+1))} = \text{constant} \ \ \forall \text{k} \in \mathbb{Z} \right\} .$$

Therefore, by using (1.19) in Theorem 1.2.2, the Haar function is defined as

$$h(x) = \begin{cases} 1 & for \ \ 0 \le x < 1/2, \\ -1 & for \ \ 1/2 \le x < 1, \\ 0 & otherwise, \end{cases}$$

and the Haar basis is formed by functions $h_{j,k}(x) = 2^{-j/2}h(2^{-j}x - k), j, k \in \mathbb{Z}$.

Historically, the first orthonormal wavelet basis is the Haar basis, constructed long before the term "wavelet" was coined. However, the Haar function is not continuous, it is not suitable for the analysis of smooth functions.

1.3 Orthonormal Bases of Compactly Supported Wavelets

In the following, we discuss how to construct compactly supported orthogonal wavelets briefly. The easiest way to ensure compact support for the wavelet ψ is to choose the scaling function ϕ with compact support. Let ϕ be defined as in (1.10) with compact support. It follows that

$$h_n = \sqrt{2} \int \phi(x)\overline{\phi(2x - n)}dx \ \ , \tag{1.21}$$

and only finitely many h_n are non-zero, so that ψ reduces to a finite linear combination of compactly supported functions (see (1.19)), and therefore automatically has compact support itself.

For compactly supported ϕ, the 2π-periodic function m_0

$$m_0(\xi) = \frac{1}{\sqrt{2}} \sum_n h_n e^{-in\xi} \ \ , \tag{1.22}$$

is a trigonometric polynomial. By the orthonormality condition (1.9) of $\{\phi_{0,n} : n \in \mathbb{Z}\}$ and the assumption that $m_0(0) = 1$, we obtains

$$|m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1 \quad \forall \xi \ . \tag{1.23}$$

The following theorem from [8] tell us that m_0 is necessary and sufficient of the form

$$m_0(\xi) = (\frac{1 + e^{-i\xi}}{2})^N \mathcal{L}(\xi) \ , \tag{1.24}$$

with $N \geq 1$, and $\mathcal{L}(\xi)$ is a trigonometric polynomial.

Theorem 1.3.1 *A trigonometric polynomial m_0 of the form*

$$m_0(\xi) = (\frac{1 + e^{-i\xi}}{2})^N \mathcal{L}(\xi)$$

satisfies (1.23) if and only if $L(\xi) = |\mathcal{L}(\xi)|^2$ can be written as

$$L(\xi) = P(\sin^2 \xi/2)$$

with

$$P(y) = P_N(y) + y^N R(\frac{1}{2} - y) \tag{1.25}$$

where

$$P_N(y) = \sum_{k=0}^{N-1} \binom{N-1+k}{k} y^k \tag{1.26}$$

and R is an odd polynomial, chosen such that $P(y) \geq 0$ for $y \in [0, 1]$. □

This theorem tells us how to construct all possible trigonometric polynomial m_0 satisfying (1.23). It is not yet clear form the theorem, however, that any such m_0 leads to an orthonormal wavelet basis. We will see that in Theorem 1.3.5 that m_0 needs to satisfy certain assumption in order to form an orthonormal wavelet basis. We first see the connection of the ψ with the frame.

Definition 1.3.2 *Let H be a separable Hilbert space, and let J be a countable index set. A family of functions $\{\phi_j\}_{j\in J}$ in H is called a frame if there exist $A > 0$, $B < \infty$ so that, for all $f \in H$,*

$$A\|f\|^2 \le \sum_{j\in J} |\langle f, \phi_j\rangle|^2 \le B\|f\|^2 \quad . \tag{1.27}$$

We call A and B the frame bounds. If $A = B$, then the frame is said to be tight. A frame is called exact if it ceases to be a frame when any one of its elements is deleted.

The following theorem can be found in [21].

Theorem 1.3.3 *Assume that m_0 is a trigonometric polynomial satisfying (1.23), with $m_0(0) = 1$, and define ϕ, ψ by*

$$\widehat{\phi}(\xi) = (2\pi)^{-1/2} \prod_{j=1}^{\infty} m_0(2^{-j}\xi) \quad , \tag{1.28}$$

$$\widehat{\psi}(\xi) = -e^{-i\xi/2}\overline{m_0(\xi/2+\pi)}\widehat{\phi}(\xi) \quad .$$

Then ϕ, ψ are compactly supported L^2-functions, satisfying

$$\phi(x) = \sqrt{2}\sum_n h_n \phi(2x-n) \quad ,$$

$$\psi(x) = \sqrt{2}\sum_n (-1)^n \overline{h_{-n+1}}\phi(2x-n) \quad ,$$

where h_n is determined by m_0 via $m_0(\xi) = \frac{1}{\sqrt{2}}\sum_n h_n e^{-in\xi}$. Moreover, $\{\psi_{j,k}\}$, $j, k \in \mathbb{Z}$ constitute a tight frame for $L^2(\mathbb{R})$ with frame constant 1.

In order to ensure the frame $\{\psi_{j,k} : j, k \in \mathbb{Z}\}$ above to be a wavelet basis. We need the following conditions due to Cohen [5].

Definition 1.3.4 *A compact set K is called congruent to $[-\pi, \pi]$ modulus 2π if*

(i) $|K| = 2\pi$;

(ii) For all $\xi \in [-\pi, \pi]$, there exists $l \in \mathbb{Z}$ so that $\xi + 2l\pi \in K$.

Theorem 1.3.5 *Under the hypothesis as in the above theorem, this tight frame is an orthonormal basis if and only if m_0 satisfies one of the following equivalent conditions:*

(i) There exists a compact set K congruent to $[-\pi, \pi]$ modulus 2π and containing a neighborhood of 0 so that

$$\inf_{k>0} \inf_{\xi \in K} |m_0(2^{-k}\xi)| > 0. \tag{1.29}$$

(ii) There is no non-trivial cycle $\{\xi_1, \cdots, \xi_n\}$ in $[-\pi, \pi]$ invariant for the operation $\xi \mapsto 2\xi$ mod (2π) such that $|m_0(\xi_j)| = 1$ for all $j = 1, \cdots, n$.

(iii) The eigenvalue 1 of the $(2N-1) \times (2N-1)$-dimensional matrix A defined by

$$A_{lk} = \sum_{n=0}^{N} h_n \overline{h_k - 2l + n} \quad , \quad -N+1 \le l, k \le N+1$$

(where we assume $h_n = 0$ for $n < 0$, $n > N$) is non-degenerate.

1.3.1 Example : The Daubechies Wavelets

The Daubechies wavelets [7] are compactly supported orthogonal wavelet. Its construction is based on Theorem 1.3.1. We let π_m denote the collection of all polynomial of degree $\le m$. By considering

$$\begin{aligned} P(z) &= \frac{1}{2}\sum_{n=0}^{N} p_n z^n = \left(\frac{1+z}{2}\right)^m S(z) \\ S(z) &\in \pi_{N-m} \ , \ \ S(1) = 1 \ \text{ and } \ S(-1) \neq 0 \ , \end{aligned} \tag{1.30}$$

for some integer $N > m$ (p_k to be determined explicitly). The normalization of $S(z)$ in (1.30) is to ensure that $P(1) = 1$ and that the factor $S(z)$ does not increase the multiplicity m of the root -1 of $P(z)$. Once $S(z)$ has been determined, the orthonormal Daubechies scaling function $\phi_m(t)$ can be computed by taking the inverse Fourier transform of the infinite product

$$\widehat{\phi}_m(\omega) = \prod_{k=1}^{\infty} P(e^{-j\omega/2^k}) \ , \tag{1.31}$$

where $P(z)$ is obtained by applying (1.30). The sequence $\{p_k\}$ can also be derived from (1.30) and the corresponding dilation equation is given by

$$\phi_m(t) = \sum_{k=0}^{N} p_k \phi_m(2t-k) \ . \tag{1.32}$$

Then, the corresponding Daubechies wavelet is given by

$$\psi_m(t) = \sum_{k=-N+1}^{1} (-1)^k p_{1-k} \phi_m(2t-k) \ . \tag{1.33}$$

The main difficulty is to determine the polynomial factor $S(z)$ of $P(z)$ in (1.30). To do this, we observe that the Fourier representation of (1.32) is given by

$$\widehat{\phi}_m(\omega) = P(z)\widehat{\phi}_m(\frac{\omega}{2}) \ , \quad z = e^{-i\omega/2} \ . \tag{1.34}$$

Since we want the scaling function $\phi_m(t)$ to be orthonormal, $P(z)$ must satisfy (1.23), this means that we need

$$|P(z)|^2 + |P(-z)|^2 = 1 \qquad \forall |z| = 1 \ . \tag{1.35}$$

To transform the condition (1.35) on $P(z)$ to $S(z)$, let us consider the change of variables

$$x := \frac{1-\cos\frac{\omega}{2}}{2} = \sin^2\left(\frac{\omega}{4}\right) \ , \tag{1.36}$$

and introduce the algebraic polynomial

$$f(x) := |S(e^{-i\omega/2})|^2 \ . \tag{1.37}$$

Then we see that (1.35) is equivalent to

$$(1-x)^m f(x) + x^m f(1-x) = 1 \qquad \forall x \in \mathbb{R} \ . \tag{1.38}$$

Consider the following polynomial introduced in Theorem 1.3.1,

$$f_0(x) = \sum_{k=0}^{m-1} \binom{m+k-1}{k} x^k \quad . \tag{1.39}$$

By applying the Euclidean algorithm, we see that $f_0(x)$ is the one and only one polynomial of degree $\leq m$ satisfies (1.38). Going back to the polynomial factor $S(z)$ in (1.30), we see that the solution $f(x) = f_0(x)$ in (1.39) becomes

$$|S(e^{-i\omega/2})|^2 = \sum_{k=0}^{m-1} \binom{m+k-1}{k} \sin^{2k}(\frac{\omega}{4}) \quad , \tag{1.40}$$

simply by putting (1.36) into (1.39) and then into (1.37). Consequently, we must solve for $S(e^{-i\omega/2})$ in (1.40).

Firstly, write (1.40) as a cosine series:

$$|S(e^{-i\omega/2})|^2 = \frac{a_0}{2} + \sum_{k=1}^{m-1} a_k \cos(\frac{k\omega}{2}) \quad . \tag{1.41}$$

where, for each $k = 0, 1, \cdots, m-1$, we have

$$a_k = (-1)^k \sum_{l=0}^{m-k-1} \frac{1}{2^{2(k+l)-1}} \binom{2(k+l)}{l} \binom{m+k-l-1}{k+l} \quad . \tag{1.42}$$

Next, consider the polynomial $z^{m-1}|S(z)|^2$. By using the Riesz lemma [8], we know that its roots are in reciprocal and complex conjugate pairs. More precisely, we have

$$\begin{aligned} z^{m-1}|S(z)|^2 &= \frac{1}{2} \sum_{l=0}^{2m-2} a_{|l-m+1|} z^l \\ &= \frac{a_{m-1}}{2} \prod_{k=1}^{K} (z - r_k)(z - \frac{1}{r_k}) \\ &\quad \times \prod_{l=1}^{L} (z - z_l)(z - \overline{z}_l)(z - \frac{1}{z_l})(z - \frac{1}{\overline{z}_l}) \quad , \end{aligned} \tag{1.43}$$

where $K + 2L = m - 1$, $r_1, \cdots, r_k \neq 0$ are real numbers and $z_1, \cdots, z_L$ are complex numbers.

We then consider

$$S_{m-1}(z) := \text{constant} \times \prod_{k=1}^{K}(z - r_k)\prod_{l=1}^{L}(z - z_l)(z - \overline{z}_l) \ , \tag{1.44}$$

where $|r_1|, \cdots, |r_K|, |z_1|, \cdots, |z_L| < 1$ and let $S(z) = S_{m-1}(z)$ with the constant is chosen so that $S(1) = 1$. This determines $S(z)$.

Example (The Daubechies wavelet of order 2): For $m = 2$ in (1.42), we have $a_0 = 4$ and $a_1 = -1$, so that by (1.43),

$$z|S_1(z)|^2 = \frac{1}{2}(-1 + 4z - z^2) = -\frac{1}{2}(z - r_1)(z - \frac{1}{r_1}) \tag{1.45}$$

with

$$r_1 = 2 - \sqrt{3} \tag{1.46}$$

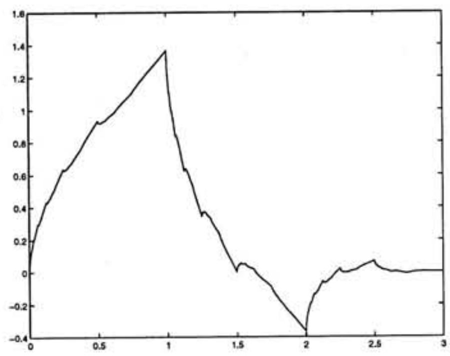
$(\frac{1}{r_1} = 2 + \sqrt{3})$. Hence, it follows from (1.44) that

$$\begin{aligned} S_1(z) &= \text{constant} \times (z - r_1) \\ &= \frac{1}{1 - r_1}(z - r_1) \\ &= \frac{1}{2}\Big\{(\sqrt{3} + 1)z - (\sqrt{3} - 1)\Big\} \ , \end{aligned} \tag{1.47}$$

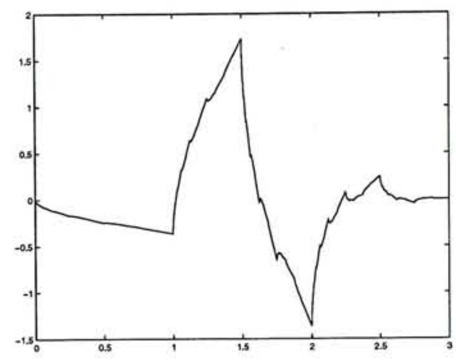
and from (1.30) that

$$\begin{aligned} P_3(z) &= \frac{1}{2}\{p_0 + p_1 z + p_2 z^2 + p_3 z^3\} = \left(\frac{1+z}{2}\right)^2 S_1(z) \\ &= \frac{1}{2}\Big\{\frac{1-\sqrt{3}}{4} + \frac{3-\sqrt{3}}{4}z + \frac{3+\sqrt{3}}{4}z^2 + \frac{1+\sqrt{3}}{4}z^3\Big\} \ . \end{aligned} \tag{1.48}$$

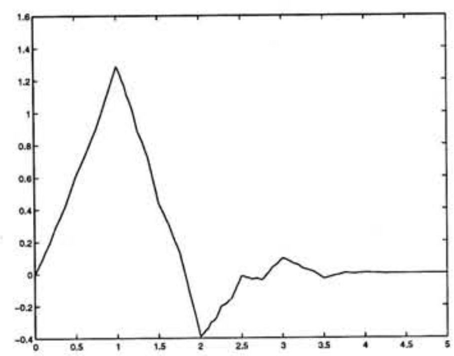
According to (1.31) and (1.33) with $N = 2m - 1 = 3$, the $P_3(z)$ in (1.48) is sufficient to determine $\phi_2(t)$ and $\psi_2(t)$. The following figures show the graphs of the Daubechies scaling functions $\phi_m(t)$ and the corresponding Daubechies wavelets $\psi_m(t)$, $m = 2, \cdots, 7$.



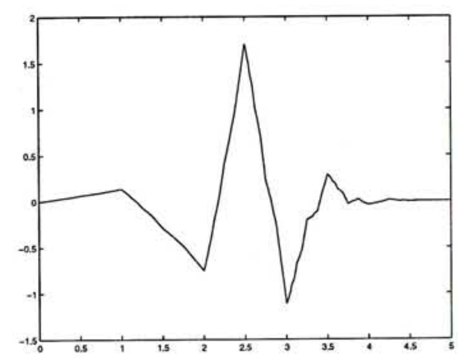
ϕ_2



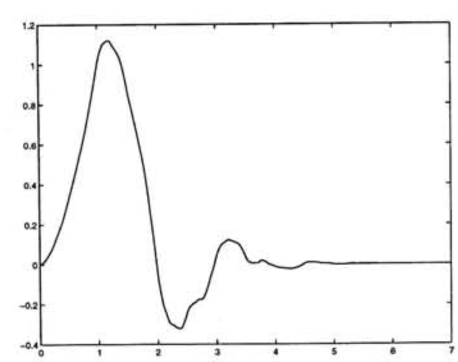
ψ_2



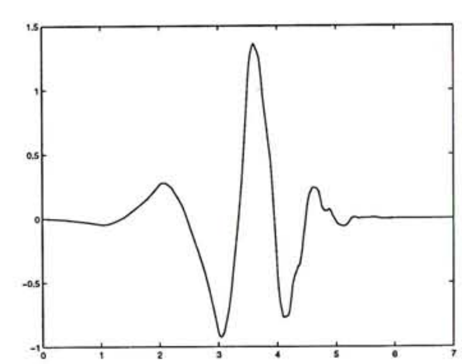
ϕ_3



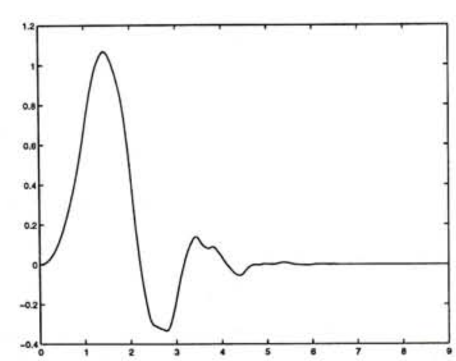
ψ_3



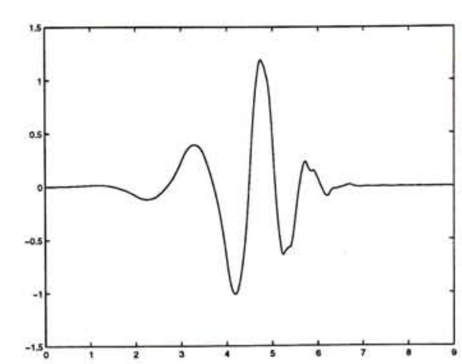
ϕ_4



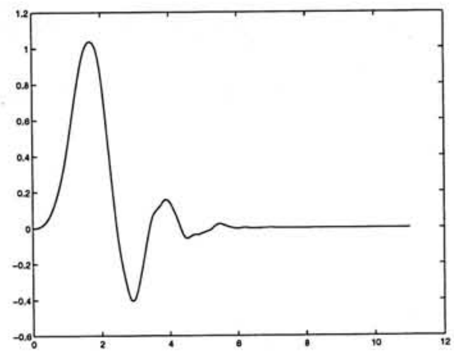
ψ_4



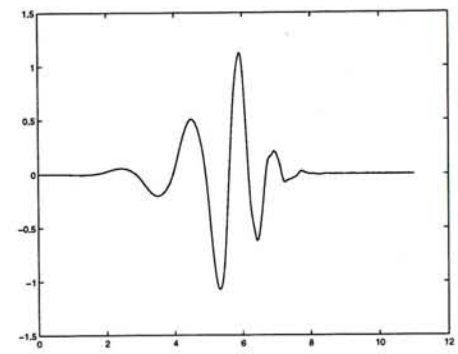
ϕ_5



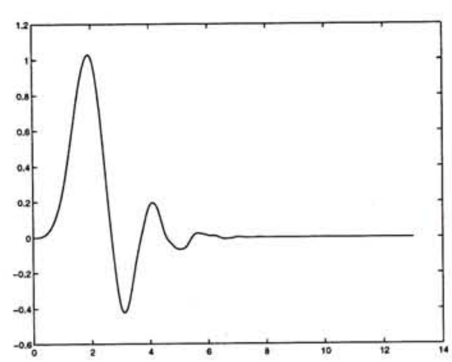
ψ_5



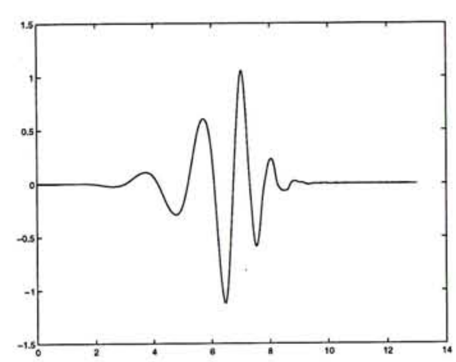
ϕ_6



ψ_6



ϕ_7



ψ_7

1.4 Wavelets in Higher Dimensions

The multivariable generalizations of one-variable wavelets can be done in many different ways. The most natural way is to use tensors to pass from one variable to several variables, i.e. functions of the form $f(x_1, \cdots, x_d) = f(x_1) \cdots f(x_d)$. We will discuss two methods in this section briefly.

1.4.1 Tensor product method

Given d functions of one variable $f^j(x)$, $j = 1, \cdots, d$, we will form the function of d variables $f^1 \otimes f^2 \otimes \cdots \otimes f^d = \otimes_{j=1}^d f^j$ defined as

$$\bigotimes_{j=1}^{d} f^j(x_1, \cdots, x_d) = \prod_{j=1}^{d} f^j(x_j) \quad . \tag{1.49}$$

If we have closed subspaces $X_j \subset L^2(\mathbb{R})$, $j = 1, \cdots, d$, we can form a closed subspace $X_1 \otimes \cdots \otimes X_d$ of $L^2(\mathbb{R}^d)$ defined as the closed linear span in $L^2(\mathbb{R}^d)$ of all functions $\otimes_{j=1}^d f^j$ where $f^j \in X_j$ for all $j = 1, \cdots, d$.

Since we know that if $\{f_s^j\}_{s\in A_j}$ are orthonormal bases in the subspaces $X_j \subset L^2(\mathbb{R})$ for some index set A_j and $j = 1 \cdots d$. Then the system

$$\left(\bigotimes_{j=1}^{d} f_{s_j}^j\right)_{(s_1,\cdots,s_d)\in A_1\times\cdots\times A_d} \tag{1.50}$$

is an orthonormal basis in $\otimes_{j=1}^d X_j$. We also know that $\otimes_{j=1}^d L^2(\mathbb{R}) = L^2(\mathbb{R}^d)$. So if we want to obtain an orthonormal basis in $L^2(\mathbb{R}^d)$, the most natural approach is to take d orthonormal bases $\{\psi_n^j\}_{n\in A_j}$ in $L^2(\mathbb{R})$ for $j = 1, \cdots, d$, and to form an orthonormal system $\otimes_{j=1}^d \psi_{n_j}^j$ index by the set $A = A_1 \times \cdots \times A_d$. It follows from what we have said above that this system is an orthonormal basis in $L^2(\mathbb{R}^d)$. When we apply this procedure to wavelet bases in $L^2(\mathbb{R})$, we obtain the following:

Proposition 1.4.1 *Let $\{\psi_j\}_{j=1}^d$ be wavelets on $\mathbb{R}$ and let*

$$\Psi(x_1, \cdots, x_d) = \prod_{j=1}^{d} \psi_j(x_j) \quad ,$$

then the system $2^{(j_1+\cdots+j_d)/2}\Psi(2^{j_1}x_1 - k_1, \cdots, 2^{j_d}x_d - k_d)$ for all $j_1, \cdots, j_d, k_1, \cdots, k_d \in \mathbb{Z}$ forms as orthonormal basis in $L^2(\mathbb{R}^d)$.

1.4.2 Multiresolution Analysis in $\mathbb{R}^d$

In order to define a Multiresolution Analysis in $\mathbb{R}^d$, we need to fix some notations first. The translation will be given by elements of $\mathbb{Z}^d$, and the dilation will be $x \mapsto Ax$. We assume that A is a fixed linear map $A : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$A(\mathbb{Z}^d) \subset \mathbb{Z}^d \quad , \tag{1.51}$$

and A is expanding, which means that all (complex) eigenvalues of A have modulus greater than 1. Note that (1.51) ensures the proper coordination between translation and dilation, and implies that the matrix A has integer entries.

We have the fact that every invertible linear map $A : \mathbb{R}^d \to \mathbb{R}^d$ induces a unitary operator on $L^2(\mathbb{R}^d)$ by the formula

$$U_A f(x) = |\det A|^{1/2} f(Ax) \quad , \tag{1.52}$$

for $f \in L^2(\mathbb{R}^d)$.

Definition 1.4.2 *A wavelet set associated with a dilation matrix A is a finite set of functions $\Psi^r(x) \in L^2(\mathbb{R}^d)$, $r = 1, 2, \cdots, s$ such that the system*

$$\{|\det A|^{j/2}\Psi^r(A^j x - \gamma)\}, \qquad r = 1, \cdots, s, \quad j \in \mathbb{Z} \quad \text{and} \quad \gamma \in \mathbb{Z}^d$$

is an orthonormal basis in $L^2(\mathbb{R}^d)$.

Analogous to the one-dimensional case we will use the following notation: for a function F (Φ, Ψ etc.) on $\mathbb{R}^d$, we let

$$F_{j,\gamma}(x) = |\det A|^{-j/2} F(A^{-j}x - \gamma),$$

where $j \in \mathbb{Z}$, $\gamma \in \mathbb{Z}^d$ and A is the dilation matrix. A multidimensional MRA in $\mathbb{R}^d$ is a direct generalization of MRA in $\mathbb{R}$. All we need to do is to replace the dilation factor 2 by the dilation matrix A. We have the following lemma which similar to the one dimensional case (see (1.13)).

Lemma 1.4.3 *Let $\{V_j\}_{j\in\mathbb{Z}}$ be a MRA with a scaling function $\Phi(x)$. A function $f \in V_{-1}$ if and only if $\widehat{f}(A^*\xi) = m_f(\xi)\widehat{\Phi}(\xi)$, where $m_f(\xi)$ is a $2\pi\mathbb{Z}^d$-periodic function and*

$$\int_{[0,2\pi]^d} |m_f(\xi)|^2 d\xi = \frac{(2\pi)^d}{|\det A|} \int_{\mathbb{R}^d} |f(x)|^2 dx.$$

Let us now fix another notation to be used in the rest of this section. Let $E_0, E_1, \cdots, E_{q-1}$ be different cosets of $A(\mathbb{Z}^d)$ in $\mathbb{Z}^d$ with $q = |\det A|$. Assume that we have a function $G \in V_{-1}$. From Lemma 1.4.3, we get a $2\pi\mathbb{Z}^d$-periodic function

$$m_G(\xi) = \sum_{\gamma\in\mathbb{Z}^d} b(\gamma) e^{i\langle\xi,\gamma\rangle}$$

such that $\widehat{G}(A^*\xi) = m_G(\xi)\widehat{\Phi}(\xi)$. Let us denote $m_G^r(\xi)$, $r = 0, \cdots, q-1$ by

$$m_G^r(\xi) = \sum_{\gamma \in E_r} b(\gamma) e^{i\langle \xi, \gamma \rangle}.$$

Clearly, $m_G(\xi) = \sum_{r=0}^{q-1} m_G^r(\xi)$. Since each coset E_r can be written as $a_r + A(\mathbb{Z}^d)$ for appropriate $a_r \in \mathbb{Z}^d$, we get

$$\begin{aligned} m_G^r(\xi) &= e^{i\langle \xi, a_r \rangle} \sum_{\gamma \in A(\mathbb{Z}^d)} \tilde{b}(\gamma) e^{i\langle \xi, \gamma \rangle} && (1.53) \\ &= e^{i\langle \xi, a_r \rangle} \sum_{\gamma \in \mathbb{Z}^d} c(\gamma) e^{i\langle A^*\xi, \gamma \rangle} \quad , && (1.54) \end{aligned}$$

for some coefficients $\tilde{b}(\gamma)$ and $c(\gamma)$ depend on the function G and E_r. Writing $\mu_G^r(\xi) = \sum_{\gamma \in \mathbb{Z}^d} c(\gamma) e^{i\langle \xi, \gamma \rangle}$, we have

$$m_G^r(\xi) = e^{i\langle \xi, a_r \rangle} \mu_G^r(A^*\xi). \tag{1.55}$$

Proposition 1.4.4 *Let $G_0, \cdots, G_{q-1} \in V_{-1}$ and denote $\mu_{G_k}^r(\xi)$ by $\mu_k^r(\xi)$ for simplicity. Then we have*

(i) The system $\{G_0(t-\gamma)\}_{\gamma \in \mathbb{Z}^d}$ is an orthonormal system if and only if

$$\sum_{r=0}^{q-1} |\mu_0^r(\xi)|^2 = 1 \quad a.e.$$

(ii) The system $\{G_j(t-\gamma)\}_{\gamma \in \mathbb{Z}^d,\ j=0,\cdots,s}$, where $0 \le s \le q-1$, is an orthonormal system if and only if vectors $v_j(\xi) = (\mu_j^r(\xi))_{r=0}^q$ are orthonormal in $\mathbb{C}^q$ for almost all $\xi \in \mathbb{R}^d$.

(iii) The system $\{G_j(t-\gamma)\}_{\gamma \in \mathbb{Z}^d,\ j=0,1,\cdots,q-1}$ is an orthonormal basis in V_{-1} if and only if the matrix

$$U(\xi) = [\mu_j^r(\xi)]_{r,j=0,1,\cdots,q-1}$$

is unitary for almost all $\xi \in \mathbb{R}^d$.

Now, let us see how to construct a wavelet set associated with a MRA by using Proposition 1.4.4. As in one dimensional case, let us write

$$W_{j+1} = V_j \ominus V_{j+1} \quad , \quad j \in \mathbb{Z} \ , \tag{1.56}$$

by using (1.7) and (1.52), we obtain

$$U_A^j(V_j) = V_0 \text{ and } U_A^j(W_j) = W_0 \quad , \quad \forall j \in \mathbb{Z} \tag{1.57}$$

and $L^2(\mathbb{R}^d) = \oplus_{j\in\mathbb{Z}} W_j$. Thus in order to find a wavelet set associated with a MRA, it suffices to find a wavelet set of functions $\Psi^s(x) \in W_0$ for $s \in C$ such that $\{\Psi^s(t-\gamma)\}_{\gamma\in\mathbb{Z}^d, s\in C}$ is an orthonormal basis in W_0. But then

$$\{\Psi^s(t-\gamma)\}_{\gamma\in\mathbb{Z}^d, s\in C} \cup \{\Phi(t-\gamma)\}_{\gamma\in\mathbb{Z}^d} \ ,$$

where Φ is a scaling functions, is an orthonormal basis in V_{-1}. Moreover, Proposition 1.4.4 tells us that such a wavelet set has to have cardinality $(q-1)$ and give us a way to construct it, We simply take the function

$$\mu_0^r = \mu_\Phi^r(\xi), \quad r = 0, 1, \cdots, q-1$$

as described in (1.53) to (1.55), and choose $2\pi\mathbb{Z}^d$-periodic functions $\mu_j^r(\xi)$ for $r = 0, 1, \cdots, q-1$ and $j = 1, 2, \cdots, q-1$ so that the matrix

$$U(\xi) = [\mu_j^r(\xi)]_{r,j=0,1,\cdots,q-1}$$

is unitary. Having done this we define functions $\Psi^s(x)$, for $s = 1, 2, \cdots, q-1$ by

$$\widehat{\Psi}^s(A^*\xi) = \sum_{r=0}^{q-1} e^{i\langle\xi, a_r\rangle} \mu_s^r(A^*\xi)\widehat{\Phi}(\xi) \ ,$$

where the a_r's are representatives of different cosets $A(\mathbb{Z}^d)$ in $\mathbb{Z}^d$. It follows form Proposition 1.4.4 (iii) that $\{\Psi^s\}_{s=1}^{q-1}$ is a wavelet set associated with the given MRA. We can summarize our discussion in the following theorem. The reader can refer to [26] for the proofs and details.

Theorem 1.4.5 *For every MRA on $\mathbb{R}^d$, associated with a dilation matrix A, there exists an associated wavelet set consisting of $q-1$ functions, where $q = |\det A|$.*

1.5 Generalization to frames

Frames were introduced by Duffin and Schaeffer in the context of non-harmonic Fourier series (i.e. expansions of functions in $L^2([0,1])$ in complex exponentials $\exp(i\lambda_n x)$, where $\lambda_n \neq 2\pi n$). Frame analysis has seen a recent resurgence with the advent of wavelet theory and the continuing development of Gabor analysis. We will see in the following that a frame in a Hilbert space H are sets of non-independent vectors and they can nevertheless be used to write a straightforward and completely explicit expansion for every vector in H, which is clearly a generalization of orthonormality. Expository treatments of frames can be found in [8].

The following proposition summarize some of the useful properties of frames that we are going to use in the next few chapters.

Proposition 1.5.1 *Let H be a separable Hilbert space and $\{f_i\}_{i\in I}$ be a frame for H with frame bounds A, B.*

(i) If $\{f_i\}_{i\in I}$ is a tight frame with frame bound $A = 1$, and if $\|f_i\| = 1$ for all $i \in I$, then $\{f_i\}_{i\in I}$ constitute an orthonormal basis.

(ii) The coefficient mapping $V : H \to l^2$ defined by $Vf = \{\langle f, f_i\rangle\}_{i\in I}$ is continuous and injective, with $\|V\|^2 \leq B$.

(iii) The adjoint $V^* : l^2 \to H$ is the continuous map defined by $V^*\{c_i\} = \sum c_i f_i$, and satisfies $\|V^*\|^2 \leq B$. In particular,

$$\forall \{c_i\} \in l^2, \quad \|\sum_i c_i f_i\|^2 \leq B \sum_i |c_i|^2 \ .$$

(iv) The frame operator $Sf = V^*Vf = \sum \langle f, f_i\rangle f_i$ is a positive, continuous, and invertible mapping of H onto itself. The frame definition is equivalent to the property $AI \leq S \leq BI$.

(v) Define $\tilde{f}_i = S^{-1} f_i$. *Then the dual frame* $\{\tilde{f}_i\}$ *is a frame for* H *with frame bounds* B^{-1}, A^{-1}.

(vi) The following series converge unconditionally in the norm of H*:*

$$\forall f \in H, \qquad f = \sum_i \langle f, \tilde{f}_i \rangle f_i = \sum_i \langle f, f_i \rangle \tilde{f}_i \quad . \tag{1.58}$$

The frame is exact if and only if equation (1.58) is the unique representation of f *as* $f = \sum c_i f_i$ *or* $f = \sum d_i \tilde{f}_i$.

(vii) If $T : H \to H$ *is a continuous bijection, then* $\{Tf_i\}_{i \in I}$ *is a frame for* H *with frame bounds* $A\|T^{-1}\|^{-2}$, $B\|T\|^2$, *frame operator* TST^*, *and dual frame* $\{(T^*)^{-1}\tilde{f}_i\}_{i \in I}$.

(viii) $\{f_i \otimes f_j\}_{(i,j) \in I \times I}$ *is a frame for* $H \otimes H$ *with frame bounds* A^2 *and* B^2, *frame operator* $S \otimes S$, *and dual frame* $\{\tilde{f}_i \otimes \tilde{f}_j\}_{(i,j) \in I \times I}$.

Chapter 2

Wavelet Bases Numerical Algorithm

Numerical algorithm using wavelet bases are similar to other transform method in that vectors and operators are expanded into a basis and computations are taken place in the new system of coordinates. This approach hopes to achieve faster computation by the recursive nature of wavelets, the well localization in both time and frequency domains, and the vanishing moments property. Wavelets bases algorithm exhibit a number of new and important properties.

This chapter surveys an algorithm of Beylkin, Coifman, and Rokhlin [1] for a rapid numerical calculations of the Calderón-Zygmund operators and pseudodifferential operators.

2.1 The Algorithm in Wavelet Bases

We have introduced the Haar basis in chapter one, and in this section we will consider a numerical method to calculate the wavelet coefficients for $f \in L^2(\mathbb{R})$. In the second part, we will also discuss the method can be modified to the wavelets with vanishing moments.

2.1.1 Definitions and Notations

Let $I = I_{j,k} = [2^{-j}(k-1), 2^{-j}k] \subset \mathbb{R}$, $j, k \in \mathbb{Z}$ be dyadic intervals and $|I|$ denote its length. We will also define $g_{j,k}(x) = 2^{-j/2} g(2^{-j}x - k + 1)$ when a function g is given.

Let $\chi_{j,k}(x)$ be the normalized characteristic function on $I_{j,k}$ given by

$$\chi_{j,k}(x) = \begin{cases} |I_{j,k}|^{-1/2} & \text{for } x \in I_{j,k} \ , \\ 0 & \text{elsewhere} \ , \end{cases}$$

Let h be the Haar function and denote $\{h_{j,k}\}_{j,k\in\mathbb{Z}}$ the Haar basis as before. Clearly, the Haar function $h_{j,k}(x)$ is supported in the dyadic interval $I_{j,k}$.

Definition 2.1.1 *Let $f \in L^2(\mathbb{R})$ and $I \subset \mathbb{R}$ be a dyadic interval, the Haar coefficient d_I of f and the average s_I of f on I are define as:*

$$d_I = \int_{\mathbb{R}} f(x) h_I(x) dx \quad \textit{and} \quad s_I = \int_{\mathbb{R}} f(x) \chi_I(x) dx \ .$$

The definition implies that

$$d_I = \frac{1}{\sqrt{2}}(s_{I'} - s_{I''}) \tag{2.1}$$

where I' and I'' are the left and the right halves of I.

Equation (2.1) give us a numerical method for calculating the Haar coefficients. Suppose we are given $N = 2^n$ samples of a function f, i.e.,

$$s_k^0 = 2^{n/2} \int_{2^{-n}(k-1)}^{2^{-n}k} f(x) dx \ , \qquad k = 1, \cdots, 2^n \ . \tag{2.2}$$

This can be view as values of scaled averages of f on interval $I_{n,k}$, which is of length 2^{-n}. We then get the Haar coefficients for the intervals via (2.1) as

$$d_k^1 = \frac{1}{\sqrt{2}}(s_{2k-1}^0 - s_{2k}^0) \ .$$

We also compute the averages

$$s_k^1 = \frac{1}{\sqrt{2}}(s_{2k-1}^0 + s_{2k}^0)$$

in the intervals of length 2^{-n+1}.

By repeating this procedure, we obtain the Haar coefficients

$$d_k^{j+1} = \frac{1}{\sqrt{2}}(s_{2k-1}^j - s_{2k}^j) , \tag{2.3}$$

and the averages

$$s_k^{j+1} = \frac{1}{\sqrt{2}}(s_{2k-1}^j + s_{2k}^j) , \tag{2.4}$$

for $j = 0, \cdots, n-1$ and $k = 1, \cdots, 2^{n-j-1}$. The above process is illustrated by the pyramid scheme

$$\begin{array}{ccccccccc} \{s_k^0\} & \longrightarrow & \{s_k^1\} & \longrightarrow & \{s_k^2\} & \longrightarrow & \{s_k^3\} & \longrightarrow & \cdots \\ & \searrow & & \searrow & & \searrow & & \searrow & \\ & & \{d_k^1\} & & \{d_k^2\} & & \{d_k^3\} & & \cdots . \end{array} \tag{2.5}$$

The evaluation of the whole set of coefficients d_I and s_I in (2.5) requires $2(N-1)$ additions and $2N$ multiplication.

Form now on, let $\{V_j\}_{j\in\mathbb{Z}}$ be a MRA on $\mathbb{R}$ corresponds to a scaling function ϕ and a wavelet ψ of compact support which satisfy the following equations:

$$\phi(x) = \sqrt{2} \sum_{k=0}^{2M-1} h_{k+1}\phi(2x-k) \tag{2.6}$$

$$\psi(x) = \sqrt{2} \sum_{k=0}^{2M-1} g_{k+1}\phi(2x-k) \tag{2.7}$$

where

$$g_k = (-1)^{k-1} h_{2M-k+1} \qquad k = 1, \cdots, 2M . \tag{2.8}$$

and

$$\int \phi(x)dx = 1 . \tag{2.9}$$

Remarks: We will also use the same notations s_k^j and d_k^j to denote the wavelet coefficient and the average of a given function $f \in L^2(\mathbb{R})$ if another wavelet is used instead of the Haar basis.

As mentioned in chapter one, there are two natural ways to construct the wavelets basis in two dimension. The first one is given by *tensor product* in such a way that

$$\psi_{j,k,j',k'}(x,y) = \psi_{j,k} \otimes \psi_{j',k'} = \psi_{j,k}(x)\psi_{j',k'}(y) \ ,$$

where $j, j', k, k' \in \mathbb{Z}$. The second basis is defined by the three classes of functions in the same scale j,

$$\psi_{j,k}(x)\psi_{j,k'}(y), \quad \psi_{j,k}(x)\phi_{j,k'}(y) \quad \text{and} \quad \phi_{j,k}(x)\psi_{j,k'}(y) \quad ,$$

where $j, k, k' \in \mathbb{Z}$.

Let $\Omega = \{(x,y) \in \mathbb{R}^2 : x = y\}$ and $K : \mathbb{R}^2 \backslash \Omega \to \mathbb{R}$ be a function which is smooth away from Ω and satisfy the following properties:

$$|K(x,y)| \le \frac{1}{|x-y|} \ , \tag{2.10}$$

$$|\partial_x^M K(x,y)| + |\partial_y^M K(x,y)| \le \frac{C_M}{|x-y|^{M+1}} \ , \tag{2.11}$$

for some $M \ge 1$ and a constant C_M depends on M.

Let T be an operator defined by $K(x,y)$, which is called the kernel function of T. More precisely

$$T(f)(x) = \int K(x,y) f(y) dy \ . \tag{2.12}$$

For convenience, all the kernel functions that we are going to consider in this chapter satisfy (2.10) and (2.11).

We expand K by the two-dimensional wavelet basis,

$$K(x,y) = \sum_{j,k,k'} \alpha^j_{k,k'} \psi_{j,k}(x)\psi_{j,k'}(y) + \sum_{j,k,k'} \beta^j_{k,k'} \psi_{j,k}(x)\phi_{j,k'}(y) + \sum_{j,k,k'} \gamma^j_{k,k'} \phi_{j,k}(x)\psi_{j,k'}(y) \ . \tag{2.13}$$

The coefficients $\alpha^j_{k,k'}$, $\beta^j_{k,k'}$ and $\gamma^j_{k,k'}$ are given by

$$\alpha^j_{k,k'} = \int\int K(x,y)\psi_{j,k}(x)\psi_{j,k'}(y) dx dy \ , \tag{2.14}$$

$$\beta^j_{k,k'} = \int\int K(x,y)\psi_{j,k}(x)\phi_{j,k'}(y) dx dy \ , \tag{2.15}$$

$$\gamma^j_{k,k'} = \int\int K(x,y)\phi_{j,k}(x)\psi_{j,k'}(y)dxdy \ . \tag{2.16}$$

We also define the matrices

$$\alpha^j = \{\alpha^j_{k,k'}\} \ , \quad \beta^j = \{\beta^j_{k,k'}\} \text{ and } \gamma^j = \{\gamma^j_{k,k'}\} \ , \tag{2.17}$$

with $k, k' = 1, 2, \cdots, 2^{n-j}$. Putting (2.13) into (2.12), we obtain

$$T(f)(x) = \sum_{j,k}\psi_{j,k}(x)\sum_{k'}\alpha^j_{k,k'}d^j_{k'} + \sum_{j,k}\psi_{j,k}(x)\sum_{k'}\beta^j_{k,k'}s^j_{k'} + \sum_{j,k}\phi_{j,k}(x)\sum_{k'}\gamma^j_{k,k'}d^j_{k'} \ . \tag{2.18}$$

2.1.2 Fast Wavelet Transform

In the following, we treat the procedure being discussed as linear transformations on $\mathbb{R}^N$, viewed as the Euclidean space of all periodic sequences with period N. This assumption is needed because wavelets do not form a basis for functions on a finite interval. Wavelet basis functions overlaps in such a way that either the interval must be extended, a periodization must be performed, or the basis functions at the interval ends must be modified.

Replacing the Haar function and the characteristic function by ψ and ϕ and assuming that the coefficients s^0_k, $k = 1, \cdots, N$ are given, we replace (2.3) and (2.4) with

$$s^j_k = \sum_{m=1}^{2M} h_m s^{j-1}_{m+2k-2} \ , \quad \text{and} \quad d^j_k = \sum_{m=1}^{2M} g_m s^{j-1}_{m+2k-2} \ , \tag{2.19}$$

where $k = 1, \cdots, 2^{n-j}$. $\{s^j_k\}_{k=1}^{2^{n-j}}$ and $\{d^j_k\}_{k=1}^{2^{n-j}}$ are viewed as periodic sequences with the period 2^{n-j}.

Equation (2.19) define an orthogonal mapping

$$O_j : \mathbb{R}^{2^{n-j+1}} \to \mathbb{R}^{2^{n-j+1}} \ ,$$

converting the coefficients s^{j-1}_k with $k = 1, 2, \cdots, 2^{n-j+1}$ into the coefficients s^j_k, d^j_k with $k = 1, 2, \cdots, 2^{n-j}$. The numerical seheme is the same as (2.5). In addition, the inverse of O_j is given by

$$s^{j-1}_{2n} = \sum_{k=1}^{M} h_{2k} s^{j}_{n-k+1} + \sum_{k=1}^{M} g_{2k} d^{j}_{n-k+1} \,, \tag{2.20}$$

$$s^{j-1}_{2n-1} = \sum_{k=1}^{M} h_{2k-1} s^{j}_{n-k+1} + \sum_{k=1}^{M} g_{2k-1} d^{j}_{n-k+1} \,. \tag{2.21}$$

Computation via (2.20) and (2.21) is illustrated by the pyramid scheme

$$\begin{array}{ccccccccc}
\{s^n_k\} & \longrightarrow & \{s^{n-1}_k\} & \longrightarrow & \{s^{n-2}_k\} & \longrightarrow & \{s^{n-3}_k\} & \cdots \\
& \nearrow & & \nearrow & & \nearrow & & \\
\{d^n_k\} & & \{d^{n-1}_k\} & & \{d^{n-2}_k\} & & \{d^{n-3}_k\} & \cdots .
\end{array}$$

Obviously, given a function of the form

$$f(x) = \sum_{k=1}^{2^{n-j}} s^j_k \phi_{n-j,k}(x) + \sum_{k=1}^{2^{n-j}} d^j_k \psi_{n-j,k}(x) \,, \tag{2.22}$$

we can rewrite f as

$$f(x) = \sum_{l=1}^{2^{n-j+1}} s^{j-1}_l \phi_{n-j+1,k}(x) \,. \tag{2.23}$$

with s^{j-1}_k, $k = 1, 2, \cdots, 2^{n-j+1}$ given by (2.20) and (2.21).

Moreover, given the coefficients s^0_k, $k = 1, 2, \cdots, N$, recursive application of (2.19) yields a numerical procedure for evaluating the coefficients s^j_k, d^j_k for all $j = 1, 2, \cdots, n$, $k = 1, 2, \cdots, 2^{n-j}$, with a cost proportional to N. Similarly, given the values d^j_k for all $j = 1, 2, \cdots, n$, $k = 1, 2, \cdots, 2^{n-j}$ and s^n_1, we can reconstruct the coefficients s^0_k, for all $k = 1, 2, \cdots, N$ by using (2.20) and (2.21) recursively for $j = n, n-1, \cdots, 0$. The cost of the latter procedure is also $O(N)$. Finally, given an expression of the form

$$f(x) = \sum_{j=0}^{n} \sum_{k=1}^{2^{n-j}} s^j_k \phi_{n-j,k}(x) + \sum_{j=0}^{n} \sum_{k=1}^{2^{n-j}} d^j_k \psi_{n-j,k}(x), \tag{2.24}$$

it costs $O(N)$ steps to evaluate all coefficients s^0_k, $k = 1, 2, \cdots, N$ by recursive application of (2.23) with $j = n, n-1, \cdots, 0$.

2.2 Wavelet-Based Quadratures

Though the Haar system leads to simple algorithms, it is not very useful in actual calculations because the decay of $\alpha^j_{k,k'}$, $\beta^j_{k,k'}$ and $\gamma^j_{k,k'}$ away from the diagonal is not sufficiently fast. To have a faster decay, it is necessary to use a wavelet basis in which the elements have several vanishing moments. Let us introduce the concept of vanishing moment of a wavelet.

Definition 2.2.1 *A function f is said to has L vanishing moments if*

$$\int f(x)x^l dx = 0 \ , \qquad l = 0, \cdots, L-1. \tag{2.25}$$

The vanishing moment property of a function f simply means that f is orthogonal to lower degree polynomials. (See [7])

Consider functions ψ and ϕ (corresponding to h and χ in section 1 of this chapter), which satisfy (2.6), (2.7), (2.8) and (2.9). The coefficients $\{h_k\}_{k=1}^{2M}$ in (2.6) are chosen so that the functions $\psi_{j,k}$, $j, k \in \mathbb{Z}$, form an orthonormal wavelet basis with compact support and has M vanishing moments. For instance, the Haar function is a particular case satisfy all this assumptions with $M = 1$, $h_1 = h_2 = 1/\sqrt{2}$, $\phi = \chi$ and $\psi = h$.

Remark: The letter "M" will be reserved for the number of vanishing moments of the wavelets ψ in the following.

In the preceding subsection, we introduce a procedure for calculating the coefficients s^j_k, d^j_k for all $j \geq 1$, $k = 1, 2, \cdots, N$, given the coefficients s^0_k for $k = 1, 2, \cdots, N$. In this subsection, we introduce a set of quadrature formulae for the efficient evaluation of the coefficients s^0_k corresponding to a function f. The simplest algorithm of this kind is obtained by further assuming that there exists an integral constant τ_M such that the scaling function ϕ satisfies the condition

$$\int \phi(x + \tau_M)x^m \, dx = 0, \quad \text{for} \quad m = 1, 2, \cdots, M-1 \ , \tag{2.26}$$

$$\int \phi(x)\, dx = 1 \ ,$$

namely, that the first $M-1$ shifted moments of ϕ are equal to zero, while its integral is equal to 1.

We recall that the definition of s_k^0 is

$$\begin{aligned} s_k^0 &= 2^{n/2} \int f(x)\phi(2^n x - k + 1)\, dx \\ &= 2^{n/2} \int f(x + 2^{-n}(k-1))\phi(2^n x)\, dx \ . \end{aligned}$$

Suppose $f \in C^K(\mathbb{R})$ with $K > M$, we can expand f into a Taylor series around $2^{-n}(k - 1 + \tau_M)$, and using (2.26), we obtain

$$s_k^0 = 2^{-n/2} f(2^{-n}(k - 1 + \tau_M)) + O(2^{-n(M+1/2)}) \ . \tag{2.27}$$

In effect, (2.27) is a one-point quadrature formula for the evaluation of s_k^0. Apply the same calculation to s_k^j with $j \geq 1$, we obtain

$$s_k^j = 2^{(-n+j)/2} f(2^{-n+j}(k - 1 + \tau_M)) + O(2^{-(n-j)(M+1/2)}) \ , \tag{2.28}$$

which turns out to be extremely useful for the rapid evaluation of the coefficients of the following compressed forms of matrices.

Table 1 contains coefficients $\{h_k\}_{k=1}^{3M}$, $M = 2, 4, 6$ for the dilation equation (2.6) of ϕ for one particular choice of the shift τ. These coefficients satisfy

$$M_l = \sum_{k=1}^{3M} h_k (k - \tau_M)^l = 0 \ , \qquad l = 1, \cdots, M-1 \ ,$$

called the $M-1$ vanishing moment of the coefficients $\{h_k\}$, where τ_M is the shift and was provided by I. Daubechies [1].

It turns out that the coefficients given in Table 1 are 50% longer than Daubechies wavelets found in [7], given the same order M. Therefore, it might be desirable to adapt the numerical scheme so that the "shorter" wavelets could be used because it will leads to faster calculation.

	k	Coefficients h_k		k	Coefficients h_k
$M = 2$	1	0.038580777747887	$M = 6$	1	-0.0016918510194918
$\tau_2 = 5$	2	-0.12696912539621	$\tau_6 = 8$	2	-0.00348787621998426
	3	-0.077161555495774		3	0.019191160680044
	4	0.60749164138568		4	0.021671094636352
	5	0.74568755893443		5	-0.098507213321468
	6	0.22658426519707		6	-0.056997424478478
				7	0.45678712217269
				8	0.78931940900416
$M = 4$	1	0.0011945726958388		9	0.38055713085151
$\tau_8 = 8$	2	-0.01284557955324		10	-0.070438748794943
	3	0.024804330519353		11	-0.056514193868065
	4	0.050023519962135		12	0.036409962612716
	5	-0.15535722285996		13	0.0087601307091635
	6	-0.071638282295294		14	-0.011194759273835
	7	0.57046500145033		15	-0.0019213354141368
	8	0.75033630585287		16	0.0020413809772660
	9	0.28061165190244		17	0.00044583039753204
	10	-0.0074103835186718		18	-0.00021625727664696
	11	-0.014611552521451			
	12	-0.0013587990591632			

Table 1

We now discuss the quadrature formulae using the Daubechies wavelets but the scaling functions ϕ do not satisfy condition (2.26).

These quadrature formulae are similar to the quadrature formula (2.28) in that they do not require explicit evaluation of the function ϕ and are completely determined by the coefficients $\{h_k\}_{k=1}^{2M}$. Our interest in these quadrature formulae stems from the fact that for a given number M of vanishing moments of the basis function, the Daubechies wavelets have the support of length $2M$ compared with $3M$ for the wavelets satisfying the condition (2.26). Since our algorithms formulae in the following depend linearly on the size of the support, using Daubechies wavelets and quadrature formulae in the following will make the algorithm approximately 50% faster.

First, we explain how to compute $\{s_k^0\}_{k=1}^{M}$ if a function f is given. For convenience, we also assume that $f \in C^K(\mathbb{R})$ with $K > M$. Since

$$\begin{aligned} s_k^0 &= 2^{n/2} \int f(x)\phi(2^n x - k + 1)\, dx \\ &= 2^{n/2} \int f(x + 2^{-n}(k-1))\phi(2^n x)\, dx \ , \end{aligned}$$

we look for the coefficients $\{c_l\}_{l=0}^{M-1}$ such that

$$s_k^0 = 2^{-n/2} \sum_{l=0}^{M-1} c_l f(l + 2^{-n}(k-1)) \ , \tag{2.29}$$

for polynomials of degree less than M. Replacing f by different polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ with $n < M$ into (2.29) and solve for a_k, we arrive at the linear algebraic system for the coefficients c_l,

$$\sum_{l=0}^{M-1} l^m c_l = \int x^m \phi(x)\, dx \ , \qquad m = 0, 1, \cdots, M-1 \ , \tag{2.30}$$

where the moments of the function $\phi(x)$ on the right hand side of (2.30) are computed in terms of the coefficients $\{h_k\}_{k=1}^{2M}$.

If we can calculate the coefficients $\{c_l\}_{l=0}^{M-1}$, we obtain a quadrature formula for computing s_k^0,

$$s_k^0 = 2^{-n/2} \sum_{l=0}^{l=M-1} c_l f(l + 2^{-n}(k-1)) + O(2^{-n(M+1/2)}) \ . \tag{2.31}$$

Hence, it remains to calculate the coefficients $\{c_l\}_{l=0}^{M-1}$ from (2.30).

In order to calculate the coefficients c_l, we can make use of the following algorithm to compute the moments of the function ϕ. Let ϕ and ψ be defined as in (2.6) and (2.7), and

$$\mathcal{M}_m = \int x^m \phi(x)\, dx \ , \qquad m = 0, \cdots, M-1 \ .$$

We will also use the notations defined in (1.13) and (1.14). Applying the operator $(i\frac{d}{d\xi})^m$ on both sides of (1.13) and setting $\xi = 0$, we obtain

$$\mathcal{M}_m = 2^{-m} \sum_{j=0}^{m} \binom{m}{j} \mathcal{M}_j \mathcal{N}_{m-j}^h \tag{2.32}$$

where

$$\mathcal{N}_l^h = 2^{-\frac{1}{2}} \sum_{k=0}^{2M-1} h_{k+1} k^l \ , \qquad l = 0, \cdots, M-1 \tag{2.33}$$

Thus, we collect the term $\mathcal{M}_m$ from (2.32) and use the fact that $\mathcal{N}_0^h = 1$, we obtain

$$\mathcal{M}_m = \frac{1}{2^m - 1} \sum_{j=0}^{m-1} \binom{m}{j} \mathcal{M}_j \mathcal{N}_{m-j}^h \ .$$

Remarks: Alternatively, using (1.28), the moments $\mathcal{M}_m$ may be obtained within the desired accuracy as a limit of recursively generated sequence of vectors, $\{\mathcal{M}_m^{(r)}\}_{m=0}^{M-1}$ for $r = 1, 2, \cdots$

$$\mathcal{M}_m^{(r+1)} = \sum_{j=0}^{m} \binom{m}{j} 2^{-j(r+1)} \mathcal{M}_{m-j}^{(r)} \mathcal{N}_j^h \ ,$$

starting with

$$\mathcal{M}_m^{(1)} = 2^{-m} \mathcal{N}_m^h \ , \qquad m = 0, \cdots, M-1 \ .$$

Each vector $\{\mathcal{M}_m^{(r)}\}_{m=0}^{m=M-1}$ represents M moments of the product in (1.28) with r terms, and the iteration converges rapidly. By using this result, we can compute s_k^0 via (2.31). Notice that in both algorithm we do not have to computed the values of the function ϕ itself.

We now derive formulae to compute the coefficients of s_k^j of smooth functions replacing the pyramid scheme (2.5).

Proposition 2.2.2 *Let s_m^j be the coefficients of a smooth function at some scale j. Then*

$$s_m^{j+1} = 2^{\frac{1}{2}} \sum_{l=1}^{M} q_l \cdot s_{2m+2l-3}^j + O(2^{-(n-j)M}) \ . \tag{2.34}$$

The coefficients $\{q_l\}_{l=1}^M$ in (2.34) are solutions of the linear algebraic system

$$\sum_{l=1}^{M} (2l-1)^m \cdot q_l = M_m \ , \qquad m = 0, \cdots, M-1 \ , \tag{2.35}$$

and where

$$M_m = 2^{-\frac{1}{2}} \sum_{k=1}^{2M} h_k k^m \ , \qquad m = 0, \cdots, M-1 \ . \tag{2.36}$$

and $\sum_{k=1}^{2M} h_k = \sqrt{2}$.

Proof: Let $m_0(\xi)$ be a trigonometric polynomial defined by the coefficients $\{h_k\}_{k=1}^{2M}$,

$$m_0(\xi) := \sum_{k=1}^{2M} h_k e^{-ik\xi} \ . \tag{2.37}$$

By applying the operator $i\frac{d}{dx}$ to (2.37) α times and evaluate at $\xi = 0$, we see that

$$M_\alpha = 2^{-\frac{1}{2}} \left(i \frac{d}{d\xi} \right)^\alpha m_0(\xi)|_{\xi=0} \ , \qquad \alpha = 0, \cdots, M-1 \ . \tag{2.38}$$

Moreover, by solving a system of linear equations, the trigonometric polynomial $m_0(\xi)$ can always be written as the product,

$$m_0(\xi) = m_1(\xi) m_2(\xi) \ , \tag{2.39}$$

where we choose m_2 to be of the form

$$m_2(\xi) = \sum_{l=1}^{M} q_l e^{-i(2l-1)\xi} , \tag{2.40}$$

and m_1 is chosen so that it satisfies

$$\left(i\frac{d}{d\xi}\right)^{\alpha} m_1(\xi)|_{\xi=0} = 0 , \quad \alpha = 1, \cdots, M-1 . \tag{2.41}$$

By differentiating (2.39), setting $\xi = 0$ and using (2.41) we arrive (2.35). Solving (2.35), we find the coefficients $\{q_l\}_{l=1}^{M}$.

Since m_1 satisfies (2.41), the convolution of s_k^j with the coefficients of m_1 reduces to the one-point quadrature formula of the type in (2.28). Thus applying m_0 reduces to applying m_2 and scaling the result by $m_0(0) = \sqrt{2}$. Clearly, there are only M coefficients of m_2 compared to $2M$ of m_0, and the particular form of m_2 was chosen so that only every second entry of s_k^j, starting with $k = 1$, is multiplied by a coefficient of m_2. This means that instead of (2.19), we have (2.34). □

2.3 The Integral Operator, Standard and Non-standard Form

In this section, we introduce a class of numerical algorithm designed for the rapid multiplication of dense matrix (or integral operators) to vectors. As is well known, there are roughly N^2 operations to multiply directly a dense $N \times N$ matrix to a vector. This simple fact causes serious difficulties in large-scale computation. Most iterative methods for the solution of systems of linear equation involves the application of the matrix of the system to a sequence of recursively generated vectors, which tends to be prohibitively expensive for large scale problem. The situation is even worse if a direct solver for the linear system is used, since such solvers normally required $O(N^3)$ operations. By using the *Non-Standard* form, we transform the integral operator T into a sparse matrix, which will enable us to perform faster calculations.

2.3.1 The Standard Form

Let $T : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ be an operator with kernel as in (2.12), the standard form of the operator T is obtained by representing the operator in the tensor product basis and by representing

$$V_j = \bigoplus_{j'>j} W_{j'} \ , \tag{2.42}$$

where W_j is the orthogonal subspace of V_j in V_{j-1}. This means that we have to consider the scalar products

$$\tau(j, j', k, k') = \langle T\psi_{j,k}, \psi_{j',k'} \rangle \ ,$$

where $j, j', k, k' \in \mathbb{Z}$. We also let

$$A_j = \{\tau(j, j, k, k')\}_{k,k'=1,2,\cdots,2^{n-j}} \ ,$$

$$B_j^{j'} = \{\tau(j', j, k', k)\}_{k'=1,2,\cdots,2^{n-j'},\ k=1,2,\cdots,2^{n-j}}$$

and

$$\Gamma_j^{j'} = \{\tau(j, j', k, k')\}_{k=1,2,\cdots,2^{n-j},\ k'=1,2,\cdots,2^{n-j'}}$$

to denote the $2^{n-j} \times 2^{n-j}$, $2^{n-j'} \times 2^{n-j}$ and $2^{n-j} \times 2^{n-j'}$ matrices respectively.

For scales j , j' with $j' > j$,

$$\begin{aligned} A_j &: W_j \to W_j \ , \\ B_j^{j'} &: W_{j'} \to W_j \ , \qquad (2.43) \\ \Gamma_j^{j'} &: W_j \to W_{j'} \ , \qquad (2.44) \end{aligned}$$

If there is the coarsest scale n, then instead of (2.42), we have

$$V_j = V_n \oplus_{j'=j+1}^{n} W_{j'} \ .$$

In this case, the operators $\{B_j^{j'}\}$ and $\{\Gamma_j^{j'}\}_{j'=j+1,\cdots,n}$ are the same as in (2.43) and (2.44). In addition, for each scale j, there are operators $\{B_j^{n+1}\}$ and $\{\Gamma_j^{n+1}\}$,

$$B_j^{n+1} : V_n \to W_j \ , \tag{2.45}$$

$$\Gamma_j^{n+1} : W_j \to V_n \ . \tag{2.46}$$

In this notation, $\Gamma_n^{n+1} = \Gamma_n$ and $B_n^{n+1} = B_n$. If the number of scales is finite and V_0 is finite dimensional, then the standard form is a representation of $T_0 = P_0 T P_0$ as

$$T_0 = \left\{A_j, \{B_j^{j'}\}_{j'=j+1}^{n}, \{\Gamma_j^{j'}\}_{j'=j+1}^{n}, B_j^{n+1}, \Gamma_j^{n+1}, T_n\right\}_{j=1,\cdots,n} \ . \tag{2.47}$$

The operators (2.47) is organized as blocks of a matrix as in Figure 1. For simplicity , we put $n = 4$ to see what the matrix looks like. We will not go into detail discussion of the standard form because the non-standard form is what we are interested. The reason is that many people would say that Calderón-Zygmund operators have almost diagonal matrices in orthonormal wavelet basis. However, Yang [41] showed that this statement is not true as stated. In contrast, the non-standard matrix representation of Calderón-zygmund operators always yields almost diagonal matrices. This will be introduce in the following section.

2.3.2 The Non-standard Form

Let $T : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ be an operator with kernel $K(x, y)$ as in (2.12). Let P_j denote the projection operator of $L^2(\mathbb{R})$ onto the closed subspace V_j given by

$$(P_j f)(x) = \sum_{k\in\mathbb{Z}} \langle f, \phi_{j,k}\rangle \phi_{j,k}(x) \ .$$

Let

$$Q_j = P_{j-1} - P_j \ , \tag{2.48}$$

we see that Q_j is a projection operator of $L^2(\mathbb{R})$ on the orthogonal subspace W_j of V_j in V_{j-1}. We expand T into the following series:

$$T = \sum_{j\in\mathbb{Z}} (Q_j T Q_j + Q_j T P_j + P_j T Q_j) \ . \tag{2.49}$$

$$\begin{pmatrix} A_1 & B_1^2 & B_1^3 & B_1^4 & B_1^5 \\ \Gamma_1^2 & A_2 & B_2^3 & B_2^4 & B_2^5 \\ \Gamma_1^3 & \Gamma_2^3 & A_3 & B_3^4 & B_3^5 \\ \Gamma_1^4 & \Gamma_2^4 & \Gamma_3^4 & A_4 & B_4^5 \\ \Gamma_1^5 & \Gamma_2^5 & \Gamma_3^5 & \Gamma_4^5 & T_4 \end{pmatrix} \begin{pmatrix} d^1 \\ d^2 \\ d^3 \\ d^4 \\ s^4 \end{pmatrix} = \begin{pmatrix} \hat{d}^1 \\ \hat{d}^2 \\ \hat{d}^3 \\ \hat{d}^4 \\ \hat{s}^4 \end{pmatrix}$$

Figure 1

This is obtained by inserting Q_j into the sum $T = \sum_{j\in\mathbb{Z}}(P_{j-1}TP_{j-1} - P_jTP_j)$. If there is the coarsest scale n, then instead of (2.49), we have

$$T = \sum_{j=-\infty}^{n} (Q_jTQ_j + Q_jTP_j + P_jTQ_j) + P_nTP_n \ . \tag{2.50}$$

Moreover, if the scale $j = 0$ is the finest scale, we define

$$T_0 = \sum_{j=1}^{n}(Q_jTQ_j + Q_jTP_j + P_jTQ_j) + P_nTP_n \ , \tag{2.51}$$

By direct computation, we see that $T_0 = P_0TP_0$.

The expansions (2.49), (2.50) and (2.51) can be view as a decomposition of the operator T into a sum of contributions from different scales. Our aim is to approximate T by T_0.

Let $A_j = Q_jTQ_j$, $B_j = Q_jTP_j$ and $\Gamma_j = P_jTQ_j$, where Q_j is defined in (2.48). Then A_j, B_j and Γ_j are operators acting on the subspaces V_j and W_j such that

$$A_j : W_j \to W_j, \quad B_j : V_j \to W_j \ \text{ and } \ \Gamma_j : W_j \to V_j \ . \tag{2.52}$$

Observed that the operator A_j in (2.52) describes interaction on the scale j independent from the other scales. Moreover, since the subspace V_j contains all other subspaces $V_{j'}$ with $j' > j$, we see that the operator B_j and Γ_j describe the interaction between the scale j and all coarser scales.

Let $T_j = P_j T P_j : V_j \to V_j$, then T_j can be represented recursively by the following 2×2 matrix

$$T_j = \begin{pmatrix} A_{j+1} & B_{j+1} \\ \Gamma_{j+1} & T_{j+1} \end{pmatrix} : W_{j+1} \oplus V_{j+1} \to W_{j+1} \oplus V_{j+1} \ . \tag{2.53}$$

The *non-standard form* introduced in Beylkin [2] is a representation of an operator T as a chain of triplets

$$T \sim \{A_j, B_j, \Gamma_j\}_{j \in \mathbb{Z}} \ . \tag{2.54}$$

If there is a coarsest scale n, then

$$T \sim \Big\{ \{A_j, B_j, \Gamma_j\}_{j \in \mathbb{Z} \,:\, j \le n} \ , \ T_n \Big\} \ ,$$

where $T_n = P_n T P_n$.

If the number of scales is finite, say $j = 1, \cdots, n$ (remember we have $N = 2^n$) and if we want to apply T to an arbitrary function f. We first discretize f into samples $s_k^0, k = 1, \cdots, 2^n$ as in (2.2), then we converted it into a vector $\tilde{f} \in \mathbb{R}^{2N-2}$ consisting of all coefficients s_k^j, d_k^j determined by the algorithm (2.5) and ordered as follows:

$$\tilde{f} = (d_1^1, d_2^1, \cdots, d_{N/2}^1, s_1^1, s_2^1, \cdots, s_{N/2}^1; d_1^2, d_2^2, \cdots, d_{N/4}^2, s_1^2, s_2^2, \cdots, s_{N/4}^2; \cdots; d_1^n, s_1^n). \tag{2.55}$$

Then we construct the matrices α^j, β^j, γ^j for $j = 1, \cdots, n$ corresponding to the operator T as in (2.17), and evaluate the vectors $\widehat{s^j} = \{\widehat{s_k^j}\}$, $\widehat{d^j} = \{\widehat{d_k^j}\}$ via the formula

$$\begin{aligned} \widehat{d^j} &= \alpha^j(d^j) + \beta^j(s^j) \\ \widehat{s^j} &= \gamma^j(d^j) \ , \end{aligned}$$

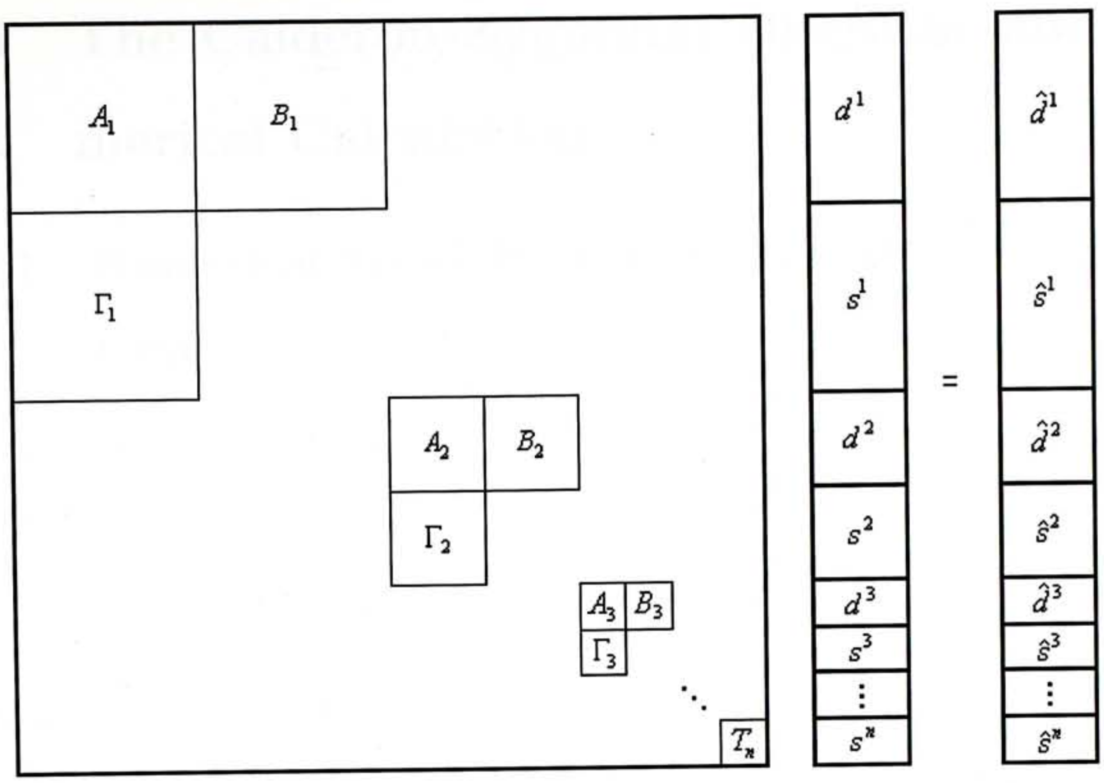


Figure 2

where $d^j = \{d^j_k\}$, $s^j = \{s^j_k\}$, $k = 1, \cdots, 2^{n-j}$, with $j = 1, \cdots, n$. Finally, we define an approximation T_0^N to T_0 by the formula

$$T_0^N(f)(x) = \sum_{j=1}^{n} \sum_{k=1}^{2^{n-j}} \left(\widehat{d}^j_k \psi_{j,k}(x) + \widehat{s}^j_k \phi_{j,k}(x) \right) , \tag{2.56}$$

which is a restriction of the operator T_0 in (2.51) on a finite dimensional subspace of $L^2(\mathbb{R})$.

The matrices α^j, β^j and γ^j, $j = 1, \cdots, n$ can be view as a single matrix and depicted in Figure 2.

The term non-standard form refers to the fact that the operator is applied to a function which is not in the usual wavelet representation. Compare this to the standard form, this non-standard representation is redundant and contains both "averages" s^j and "differences" d^j, while in the case of the standard form, the representation contains d^j's only. Such embedding into higher dimension is the price for splitting the interaction between scales. We illustrate the non-standard representation in Figure 2.

2.4 The Calderón-Zygmund Operator and Numerical Calculation

2.4.1 Numerical Algorithm to Construct the Non-Standard Form

To compute the matrix representations (2.17) of the operators A_j, B_j, Γ_j and T_j, we define an additional sets of coefficients $s^j_{k,k'}$ by

$$s^j_{k,k'} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K(x,y)\phi_{j,k}(x)\phi_{j,k'}(y)dxdy. \tag{2.57}$$

Now, given a set of coefficients $s^0_{k,k'}$ with $k, k' = 1, 2, \cdots, N$, repeated application of (2.19) to (2.14)-(2.16) to produce

$$\alpha^j_{k,k'} = \sum_{p,q=1}^{2M} g_p g_q s^{j-1}_{p+2k-2,q+2k'-2}, \tag{2.58}$$

$$\beta^j_{k,k'} = \sum_{p,q=1}^{2M} g_p h_q s^{j-1}_{p+2k-2,q+2k'-2}, \tag{2.59}$$

$$\gamma^j_{k,k'} = \sum_{p,q=1}^{2M} h_p g_q s^{j-1}_{p+2k-2,q+2k'-2}, \tag{2.60}$$

$$s^j_{k,k'} = \sum_{p,q=1}^{2M} h_p h_q s^{j-1}_{p+2k-2,q+2k'-2}, \tag{2.61}$$

with $k, k' = 1, \cdots, 2^{n-j}$, $j = 1, 2, \cdots, n$. Clearly, (2.58)-(2.61) are a two-dimensional version of the pyramid scheme (2.5) and provide an order N^2 scheme for the evaluation of the elements of all matrices α^j, β^j and γ^j with $j = 1, 2, \cdots, n$.

2.4.2 Numerical Calculation and Compression of Operators

Let $I = I_{j,k}$ and $I' = I_{j,k'}$ denoting the supports of the basis functions and label the coefficients $\alpha^j_{k,k'}$, $\beta^j_{k,k'}$ and $\gamma^j_{k,k'}$ in (2.14)-(2.16) by $\alpha_{II'}$, $\beta_{II'}$ and $\gamma_{II'}$

respectively. If the kernel $K(x,y)$ is smooth on the square $I \times I'$, then expanding $K(x,y)$ into a Taylor series around the center of the square and combining (2.10), (2.11) and remembering that the functions ψ_I and $\psi_{I'}$ are supported on the intervals I and I', we obtain the estimate

$$\begin{aligned} & |\alpha_{II'}| + |\beta_{II'}| + |\gamma_{II'}| \\ \leq \ & C \cdot |I|^{M+1} \sup_{(x,y)\in I\times I'} \left(|\partial_x^M K(x,y)| + |\partial_y^M K(x,y)| \right) , \end{aligned} \tag{2.62}$$

where M is the number of vanishing moments of the wavelet ψ. The right-hand side of (2.62) is small whenever either $|I|$ or the derivative involved are small. We use this fact to compress matrices of the integral operators by converting them to non-standard form and discarding the coefficient that are smaller than a chosen threshold.

Proposition 2.4.1 *If the wavelet basis has M vanishing moments, then for any kernel $K(x,y)$ satisfying the condition (2.10) and (2.11), the coefficients $\alpha^j_{k,k'}$, $\beta^j_{k,k'}$, $\gamma^j_{k,k'}$ in the non-standard form (see (2.14) - (2.16) and Figure 2) satisfy the estimates*

$$|\alpha^j_{k,k'}| + |\beta^j_{k,k'}| + |\gamma^j_{k,k'}| \leq \frac{C_M}{1 + |k-k'|^{M+1}} ,$$

for all $|k-k'| \geq 2M$.

Indeed, consider the operator (2.12) given by kernel function $K(x,y)$ which satisfy the estimates (2.10) and (2.11) for some $M \geq 1$. To illustrate the use of the estimates (2.10) and (2.11) for the compression of operators, we let $M = 1$ in (2.11) and consider

$$\beta_{II'} = \int \int K(x,y) h_I(x) \chi_{I'}(y) dx dy ,$$

where we assume that the distance between I and I' is greater than $|I|$. Since $\int h_I(x)dx = 0$, for all dyadic intervals $I \subset \mathbb{R}$, we have

$$\begin{aligned} |\beta_{II'}| \ & \leq \ \left| \int \int [K(x,y) - K(x_I, y)] \, h_I(x) \chi_{I'}(y) dx \, dy \right| \\ & \leq \ 2C_M \frac{|I|^2}{|x_I - y_{I'}|^2} , \end{aligned} \tag{2.63}$$

where x_I and $y_{I'}$ denote the centers of the intervals I and I' respectively. In other words, the coefficients $\beta_{II'}$ decay quadratically as a function of the distance between the intervals I, I' and for sufficient large N and finite precision of calculations, most of the entries can be discarded, leaving only a band around the diagonal. By the same argument, we can obtain similar decay properties for $\alpha_{II'}$ and $\gamma_{II'}$.

However, algorithm using the above estimates (with $M = 1$) tend to be quite inefficient, due to the slow decay of the matrix elements with their distance from the diagonal. This means that if we use a wavelet ψ with more vanishing moments, we would obtain a faster decay of the entries away from the diagonal of the non-standard form.

Note that the estimates (2.10) and (2.11) are not sufficient to conclude that $\alpha^j_{k,k'}$, $\beta^j_{k,k'}$, $\gamma^j_{k,k'}$ are bounded for $|k-k'| \leq 2M$ (For example $K(x,y) = 1/|x-y|$). If we want to strengthen the above proposition, we need to impose an extra condition, namely, the weak cancellation condition,

$$\left| \int_{I\times I} K(x,y)dxdy \right| \leq C|I| \tag{2.64}$$

for all dyadic intervals $I \subset \mathbb{R}$.

Proposition 2.4.2 *Under (2.64) and the assumptions in Proposition 2.4.1, we have*

$$|\alpha^j_{k,k'}| + |\beta^j_{k,k'}| + |\gamma^j_{k,k'}| \leq \frac{C_M}{1+|k-k'|^{M+1}}$$

for all k, k'.

Remarks: The necessary and sufficient conditions for the boundedness of T defined in (2.12), or, equivalently, for the uniform boundedness of its discretization T_0 defined in (2.51), is in fact a reformulation of the $T(1)$-Theorem of David and Journé. And we will discuss the $T(1)$-Theorem in chapter 3 in detail.

Let the operator $T_0^{N,B}$ obtained from T_0^N by setting to zero all coefficients of matrices α^j, β^j, γ^j outside of bands of width $B \geq 2M$ around their diagonals. Suppose now we approximate the operator T_0^N (2.56), in the non-standard representation of T. By using Proposition 2.4.2, we see that

$$\|T_0^{N,B} - T_0^N\| \leq \frac{C}{B^M} \log_2 N, \tag{2.65}$$

where $N = 2^n$ and C is a constant determined by the kernel K. In other words, the matrices α^j, β^j, γ^j can be approximated by banded matrices $\alpha^{j,B}$, $\beta^{j,B}$, $\gamma^{j,B}$ respectively, and the accuracy of the approximation is

$$\frac{C}{B^M} \log_2 N \ .$$

In most numerical applications , the accuracy ϵ of calculations is fixed, then B has to be chosen such that

$$\|T_0^{N,B} - T_0^N\| \leq \frac{C}{B^M} \log_2 N \leq \epsilon,$$

or, equivalently,

$$B \geq \left(\frac{C}{\epsilon} \log_2 N\right)^{\frac{1}{M}} .$$

In other words, T_0^N has been approximated to precision ϵ with its truncated version, which can be applied to arbitrary vectors for a cost proportional to $N((C/\epsilon)\log_2 N)^{1/M}$. A more detailed investigation [1] permits the estimate (2.65) to be replaced with the estimate

$$\|T_0^B - T_0\| \leq \frac{C}{B^M},$$

which making the application of the operator T_0 to an arbitrary vector with arbitrary fixed accuracy into a procedure of order exactly $O(N)$. The idea is again arising from the proof of the $T(1)$-Theorem in the next chapter.

2.5 Differential Operators in Wavelet Bases

As an example, we compute the non-standard form of differential operator $\frac{d}{dx}$ by solving a small system of linear algebraic equations.

For $l \in \mathbb{Z}$, we define

$$r_l = \int \phi(x-l)\phi'(x)\, dx \ , \tag{2.66}$$

$$\alpha_l = \int \psi(x-l)\psi'(x)\, dx \ ,$$

$$\beta_l = \int \psi(x-l)\phi'(x)\, dx \ ,$$

$$\gamma_l = \int \phi(x-l)\psi'(x)\, dx \ .$$

The matrix elements $\alpha^j_{k,k'}$, $\beta^j_{k,k'}$ and $\gamma^j_{k,k'}$ of A_j, B_j and Γ_j for the operator $\frac{d}{dx}$ are computed as follows:

$$\begin{aligned}
\alpha^j_{k,k'} &= 2^{-j}\int \psi(2^{-j}x-k)\psi'(2^{-j}x-k')2^{-j}\, dx = 2^{-j}\alpha_{k-k'} \ , \\
\beta^j_{k,k'} &= 2^{-j}\int \psi(2^{-j}x-k)\phi'(2^{-j}x-k')2^{-j}\, dx = 2^{-j}\beta_{k-k'} \ , \\
\gamma^j_{k,k'} &= 2^{-j}\int \phi(2^{-j}x-k)\psi'(2^{-j}x-k')2^{-j}\, dx = 2^{-j}\gamma_{k-k'} \ .
\end{aligned} \tag{2.67}$$

By using (2.6) to (2.9), we arrive at

$$\alpha_i = 2\sum_{k=0}^{2M-1}\sum_{k'=0}^{2M-1} g_{k+1}g_{k'+1}r_{2i+k-k'} \ ,$$

$$\beta_i = 2\sum_{k=0}^{2M-1}\sum_{k'=0}^{2M-1} g_{k+1}h_{k'+1}r_{2i+k-k'} \ ,$$

$$\gamma_i = 2\sum_{k=0}^{2M-1}\sum_{k'=0}^{2M-1} h_{k+1}g_{k'+1}r_{2i+k-k'} \ . \tag{2.68}$$

Therefore, the representation of d/dx is completely determined by the coefficients r_l in (2.66) , in other words, by the representation of d/dx on the subspace V_0.

The following proposition reduces the computation of the coefficients r_l to solving a system of linear algebraic equations.

Proposition 2.5.1

(i) If the integrals in (2.66) exists, then the coefficients r_l, $l \in \mathbb{Z}$ in (2.66) satisfy the following system of linear algebraic equations

$$r_l = 2\left[r_{2l} + \frac{1}{2}\sum_{k=1}^{M} a_{2k-1}(r_{2l-2k+1} + r_{2l+2k-1})\right] , \tag{2.69}$$

and

$$\sum_l l r_l = -1 , \tag{2.70}$$

where

$$a_{2k-1} = 2\sum_{j=0}^{M-2k} h_{j+1}h_{j+2k} , \quad k = 1, \cdots, M . \tag{2.71}$$

(ii) If $M \geq 2$, then (2.69) and (2.70) have a unique solution with a finite number of non-zero r_l, namely, $r_l \neq 0$ for $-2M+2 \leq l \leq 2M-2$ and

$$r_l = -r_{-l} \tag{2.72}$$

Proof (i): First we note that in our case, the trigonometric polynomial $m_0(\xi) = \frac{1}{\sqrt{2}}\sum_{k=0}^{2M-1} h_{k+1}e^{-i\xi k}$. Therefore

$$|m_0(\xi)|^2 = \frac{1}{2} + \frac{1}{2}\sum_{n=1}^{2M-1} a_n \cos n\xi , \tag{2.73}$$

and

$$|m_0(\xi+\pi)|^2 = \frac{1}{2} - \frac{1}{2}\sum_{k=1}^{M} a_{2k-1}\cos(2k-1)\xi + \frac{1}{2}\sum_{k=1}^{M-1} a_{2k}\cos 2k\xi , \tag{2.74}$$

where a_n are given by

$$a_n = 2\sum_{j=0}^{2M-1-n} h_{i+1}h_{i+1+n} , \quad n = 1, \cdots, 2M-1. \tag{2.75}$$

Combining (2.73) and (2.74) to satisfy (1.23) we obtain

$$\sum_{k=1}^{M-1} a_{2k}\cos 2k\xi = 0$$

and hence

$$a_{2k} = 0 \ , \qquad k = 1, \cdots, M-1 \ . \tag{2.76}$$

Therefore, we obtain (2.71).

Using (2.6) for both $\phi(x-l)$ and $\phi'(x)$ in (2.66), we obtain

$$\begin{aligned}
r_l &= \int \phi(x-l)\phi'(x)\,dx \\
&= \int \Big(\sqrt{2}\sum_{k=0}^{2M-1} h_{k+1}\phi(2x-2l-k)\Big)\Big(\sqrt{2}\sum_{k'=0}^{2M-1} h_{k'+1}\cdot 2\cdot \phi'(2x-k')\Big)\,dx \\
&= 2\sum_{k=0}^{2M-1}\sum_{k'=0}^{2M-1} h_{k+1}h_{k'+1}\int \phi(x-2l-k-k')\phi'(x)\,dx \\
&= 2\sum_{k=0}^{2M-1}\sum_{k'=0}^{2M-1} h_{k+1}h_{k'+1}r_{2l+k-k'} .
\end{aligned}$$

Substituting $m = k - k'$, we obtain

$$r_l = 2\sum_{k=0}^{2M-1}\sum_{m=k}^{k-2M+1} h_{k+1}h_{k-m+1}r_{2l+m} \ .$$

Changing the order of summation and using $\sum_{k=0}^{2M-1} h_k^2 = 1$, we arrive at

$$r_l = 2r_{2l} + \sum_{n=1}^{2M-1} a_n(r_{2l-n} + r_{2l+n}), \qquad l \in \mathbb{Z} \ , \tag{2.77}$$

where a_n are given in (2.75). Using (2.76), we obtain (2.69) from (2.77).

In order to obtain (2.70), we first consider $\phi(x-y)$ as a function of y and expand $\widehat{\phi}(\xi)$ into a Fourier series, we then differentiate $\widehat{\phi}$ with respect to ξ and evaluate it at $\xi = 0$, we finally obtain,

$$\sum_{l\in\mathbb{Z}} l\cdot\phi(x-l) = x - \int s\phi(s)\,ds \ .$$

Using (2.9), (2.66) and the above equation, we obtain (2.70).

(ii): Let's recall a fact from [7]: if $M \geq 2$, then

$$|\widehat{\phi}(\xi)|^2|\xi| \leq C(1+|\xi|)^{-1-\epsilon} \ ,$$

where $\epsilon > 0$, and hence the integral (2.66) is absolutely convergent. The existence of the solution of the system of equations (2.69) follows from the existence of the integral in (2.66). Since the scaling function ϕ has compact support, there are only finite number of non-zero coefficients r_l. The specific interval $-2M+2 \leq l \leq 2M-2$ is obtained by the direct examination of (2.66).

Let us show that

$$\sum_l r_l = 0 \; . \tag{2.78}$$

Multiplying (2.69) by $e^{-il\xi}$ and summing over l, we obtain

$$\widehat{r}(\xi) = 2\Big(\widehat{r}_{even}(\xi/2) + \widehat{r}_{odd}(\xi/2) \sum_{k=1}^{M} a_{2k-1} \cos\,((2k-1)\frac{\xi}{2})\Big) \; , \tag{2.79}$$

where

$$\widehat{r}(\xi) = \sum_l r_l e^{-il\xi} \; ,$$

$$\widehat{r}_{even}(\xi/2) = \sum_l r_{2l} e^{-il\xi} \; ,$$

and

$$\widehat{r}_{odd}(\xi/2) = \sum_l r_{2l+1} e^{-i(2l+1)\xi/2} \; .$$

Noticing that

$$2\widehat{r}_{even}(\xi/2) = \widehat{r}(\xi/2) + \widehat{r}(\xi/2+\pi)$$

and

$$2\widehat{r}_{odd}(\xi/2) = \widehat{r}(\xi/2) - \widehat{r}(\xi/2+\pi) \; ,$$

using (2.73), we obtain from (2.79)

$$\widehat{r}(\xi) = \Big(\widehat{r}(\xi/2) + \widehat{r}(\xi/2+\pi) + (2|m_0(\xi/2)|^2 - 1)(\widehat{r}(\xi/2) - \widehat{r}(\xi/2+\pi))\Big) \; .$$

Finally, using (1.23), we arrive at

$$\widehat{r}(\xi) = 2\Big(|m_0(\xi/2)|^2\, \widehat{r}(\xi/2) + |m_0(\xi/2+\pi)|^2\, \widehat{r}(\xi/2+\pi)\Big) \; .$$

Setting $\xi = 0$ in above, we obtain $\widehat{r}(0) = 2\widehat{r}(0)$ and thus (2.78).

To prove the uniqueness of the solution r_l of (2.69) and (2.70), we consider the operator T_j defined by these coefficients on the subspace V_j and apply it to a sufficiently smooth function f, i.e.

$$(T_j f)(x) \sum_{k\in\mathbb{Z}} \Big(2^{-j} \sum_{l} r_l f_{j,k-l}\Big) \phi_{j,k}(x) \ , \tag{2.80}$$

where

$$f_{j,k-l} = 2^{-j/2} \int_{-\infty}^{\infty} f(x)\phi(2^{-j}x - k + l)dx \ . \tag{2.81}$$

Rewriting (2.81)

$$f_{j,k-l} = 2^{-j/2} \int_{-\infty}^{\infty} f(x - 2^j l)\phi(2^{-j}x - k)dx \ ,$$

and expanding $f(x - 2^j l)$ in the Taylor series at the point x we obtain

$$f_{j,k-l} = \int_{-\infty}^{\infty} f(x)\phi_{j,k}(x)dx - 2^j l \int_{-\infty}^{\infty} f'(x)\phi_{j,k}(x)dx + 2^{2j}\frac{l^2}{2} \int_{-\infty}^{\infty} f''(\widetilde{x})\phi_{j,k}(x)dx \ , \tag{2.82}$$

where $\widetilde{x} = \widetilde{x}(x, x - 2^j l)$ and $|\widetilde{x} - x| \leq 2^j l$. Substituting (2.82) into (2.80), and using (2.78) and (2.70) we obtain

$$\begin{aligned} (T_j f)(x) &= 2^j \sum_{k\in\mathbb{Z}} \Big(\frac{1}{2} \sum_l r_l l^2 \int_{-\infty}^{\infty} f''(\widetilde{x})\phi_{j,k}(x)dx\Big) \phi_{j,k}(x) \\ &\quad + \sum_{k\in\mathbb{Z}} \Big(\int_{-\infty}^{\infty} f'(x)\phi_{j,k}(x)dx\Big) \phi_{j,k}(x) \ . \end{aligned}$$

Form the above equation, it is clear that as $j \to -\infty$, operators T_j and d/dx coincide on smooth function. Since smooth functions are dense in L^2, $T_{-\infty} = d/dx$ and, hence, the solution to (2.69) and (2.70) is unique. Since r_l can be written as $r_l = -\int |\widehat{\phi}|^2(i\xi)e^{-il\xi}\, d\xi$, then (2.72) follows immediately. $\Box$.

We list the $\{r_l\}$ in Proposition 2.5.1 in Table 2 using the Daubechies wavelets with $M = 2, 3, 4, 5, 6$. Once the coefficients $\{r_l\}$ are obtained, the matrix representation of d/dx is completely determined by using (2.67) and (2.68).

	l	Coefficients r_l		l	Coefficients r_l
$M=2$	1	-0.66666666	$M=6$	1	-0.850136661
	2	0.833333333		2	0.258552944
				3	$-7.2440589e^{-2}$
$M=3$	1	-0.74520547		4	$1.4545511e^{-2}$
	2	0.14520579		5	$-1.5885615e^{-3}$
	3	0.01461187		6	$4.29689157e^{-6}$
	4	0.00034246		7	$1.202657519e^{-5}$
				8	$4.20691204e^{-7}$
$M=4$	1	-0.7930095		9	$-2.899666805e^{-9}$
	2	0.1919989		10	$6.96865115e^{-13}$
	3	- 0.0335802			
	4	0.00222404			
	5	0.0001722			
	6	$8.408505e^{-7}$			
$M=5$	1	-0.82590601			
	2	0.228820187			
	3	$-5.335257e^{-2}$			
	4	$7.461396e^{-3}$			
	5	$-2.3923582e^{-4}$			
	6	$-5.40473016e^{-5}$			
	7	$-2.52411714e^{-7}$			
	8	$-2.6960479e^{-10}$			

Table 2

Chapter 3

$T(1)$-Theorem of David and Journé

The original proof of the $T(1)$-Theorem of Guy David and Jean-Lin Journé [9] require the well-known near-orthogonality lemma of Cotlar and Stein. In the first section, we survey the result of Y. Meyer to see how to use wavelet theory to simplify the proof of the $T(1)$-Theorem. In section 2, we discuss some recent results on the theorem with a weaker condition on the distribution kernel.

3.1 Definitions and Notations

In this chapter, we fix a MRA $\{V_j\}_{j\in\mathbb{Z}}$ of $L^2(\mathbb{R}^n)$ with a C^1-scaling function ϕ, which has compact support, and $\{\phi(x-k) : k \in \mathbb{Z}^n\}$ is an orthonormal basis of V_0. Let $\{\psi_\epsilon : \epsilon \in \{0,1\}^n \backslash \{(0,\cdots,0)\}\}$ be the corresponding wavelet set. Let $\psi_\lambda(x) = 2^{nj/2}\psi_\epsilon(2^j x - k)$, where $\lambda = 2^{-j}k + 2^{-j-1}\epsilon$, $j \in \mathbb{Z}$, $k \in \mathbb{Z}^n$ and $\epsilon \in \{0,1\}^n\backslash\{(0,\cdots,0)\}$. We let Λ denote the set of all such λ. We further let $Q(\lambda) = \{x \in \mathbb{R}^n : 2^j x - k \in [0,1)^n\}$ be the dyadic cube with measure 2^{-nj} such that the support of ψ_λ is contained in $mQ(\lambda)$, for some fixed integer $m \geq 1$ and $mQ(\lambda)$ is defined by $2^j x - k \in [-m/2 + 1/2, m/2 + 1/2)^n$.

3.1.1 $T(1)$ Operator

Let $\mathcal{D}(\mathbb{R}^n)$ be the vector space of all $\phi \in C^\infty(\mathbb{R}^n)$ whose support is compact and let $\mathcal{D}'(\mathbb{R}^n)$ be the dual space of $\mathcal{D}(\mathbb{R}^n)$. The elements in $\mathcal{D}(\mathbb{R}^n)$ are called test functions and the elements in $\mathcal{D}'(\mathbb{R}^n)$ are called distributions.

Definition 3.1.1 *Let $q \geq 1$ be an integer. For each ball $B = B(x_0, r) \subset \mathbb{R}^n$ and for every function $f \in C_c^p(\mathbb{R}^n)$ with $p \geq q$, whose support lies in B, we put*

$$N_q^B(f) = r^{n/2} \sum_{|\alpha| \leq q} r^{|\alpha|} \|\partial^\alpha f\|_\infty \quad . \tag{3.1}$$

Definition 3.1.2 *Let V be a topological vector space such that*

$$\mathcal{D}(\mathbb{R}^n) \subset V \subset L^2(\mathbb{R}^n) \ , \tag{3.2}$$

where the inclusions are continuous. Let V' be the dual space of V, we say that a continuous linear operator $T : V \to V'$ is weakly continuous on $L^2(\mathbb{R}^n)$ if there exists a constant C and an integer q such that, for every ball $B \subset \mathbb{R}^n$ and every pair of functions f and g in V with supports in B, we have

$$|\langle Tf, g \rangle| \leq C N_q^B(f) N_q^B(g) \quad . \tag{3.3}$$

Remark: We will also call elements in V and V' the test functions and distributions respectively. The assumption (3.2) implies that $L^2(\mathbb{R}^n) \subset V' \subset \mathcal{D}'(\mathbb{R}^n)$.

Definition 3.1.3 *A continuous linear operator $T : V \to V'$ is said to correspond to a singular integral operator if there are $\gamma \in (0, 1]$, C_0, $C_1 > 0$ and a kernel function $K : \Omega \to \mathbb{C}$, where $\Omega = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : x \neq y\}$, such that*

(i) $|K(x,y)| \leq C_0 |x-y|^{-n} \quad ,$

(ii) $|K(x',y) - K(x,y)| \leq C_1 |x'-x|^\gamma |x-y|^{-n-\gamma} \quad , \quad if \quad |x'-x| \leq |x-y|/2 \quad ,$

(iii) $|K(x,y') - K(x,y)| \leq C_1 |y'-y|^\gamma |x-y|^{-n-\gamma} \quad , \quad if \quad |y'-y| \leq |x-y|/2 \quad ,$

(iv) $Tf(x) = \int K(x,y)f(y)dy$*, for every function* $f \in V$ *and every* x *not in support of* f*.*

Examples: The trivial example of kernel satisfies (i), (ii), (iii) of Definition 3.1.3 is of course $K(x,y) = 1/|x-y|^n$. Another example is given by the following function L defined on $\mathbb{R}^n \backslash \{0\}$: We start with a continuous function $\Omega(x)$ on $\mathbb{R}^n \backslash \{0\}$ which is homogeneous of degree 0, i.e., $\Omega(rx) = \Omega(x)$ for all $r > 0$, and which satisfies the condition $\int_{S^{n-1}} \Omega(x) d\sigma(x) = 0$, where $d\sigma$ is the rotation-invariant probability measure on the unit sphere S^{n-1}. We assume that $\Omega(x)$ satisfies

$$|\Omega(x) - \Omega(x')| \le C|x - x'|^\gamma , \qquad \forall\, x, x' \in S^{n-1} .$$

If we let $L(x) = \Omega(x)/|x|^n$ and define $K(x,y) = L(x-y)$, then $K(x,y)$ is the kernel satisfy (i), (ii), (iii) in the Definition 3.1.3.

Definition 3.1.4 *Let* T *be an operator as in the Definition 3.1.3.* T *is said to be a Calderón-Zygmund operator if* T *can be extended to a bounded linear operator on* $L^2(\mathbb{R}^n)$*.*

Remark: If T is a Calderón-Zygmund operator, the kernel function of T satisfies (i), (ii), (iii), (iv) in Definition 3.1.3, which automatically satisfies the kernel condition (2.10) in Chapter 2. So we may apply the numerical algorithm introduced in Chapter 2 to calculate the operator T.

Let $T : V \to V'$ be a continuous linear operator corresponding to a singular integral as in Definition 3.1.3 with kernel function $K(x,y)$. The adjoint $T^* : V \to V'$ of T is an integral operator corresponding to the kernel function $\overline{K(y,x)}$ defined by the following relation

$$\langle T^* f, g\rangle := \langle f, Tg\rangle = \int f \cdot \overline{Tg} \ .$$

In particular, if the kernel $K(x,y)$ of T is a real-valued function, we denote $T^* = T^t$ and call it the transpose of T. In this case, the kernel function of T^t is $K(y,x)$.

We now come to the definition of $T(1)$. The difficulty is that 1 is not a test function and $T(1)$ is not well defined in the above sense. However, it can be defined based on the following lemma.

Lemma 3.1.5 *Let S be a distribution, suppose that there exists $R > 0$ such that the restriction of S to the open set $|x| > R$ is a continuous function and such that $S(x) = O(|x|^{-n-\gamma})$ as $|x| \to \infty$. If $\gamma > 0$, then the integral $\int_{\mathbb{R}^n} S(x)\, dx = \langle S, 1\rangle$ converges.*

Proof: We write $1 = \phi_0(x) + \phi_1(x)$, where $\phi_0 \in \mathcal{D}(\mathbb{R}^n)$ and $\phi_0 = 1$ in a bounded neighborhood of $|x| \leq R$. Then $\langle S, 1\rangle$ is defined by

$$\langle S, 1\rangle := \langle S, \phi_0\rangle + \int S(x)\phi_1(x)\, dx \quad .$$

Note that the integral converges absolutely. It also follows immediately that $\langle S, 1\rangle$ is independent of the decomposition. □

In the following, we will consider the real-valued kernel $K(x,y)$ only, the complex case follows immediately if the real case is done.

let $\mathcal{D}_0(\mathbb{R}^n)$ be the subspace of $f \in \mathcal{D}(\mathbb{R}^n)$ such that $\int f(x)dx = 0$. If $f \in \mathcal{D}_0(\mathbb{R}^n)$ and the support of f is contained in $|x| \leq R$, using (ii) in Definition 3.1.3,

$$T^t f(x) = \int \Big(K(y,x) - K(0,x)\Big) f(y)\, dy = O(|x|^{-n-\gamma}) \quad , \quad \text{for } |x| > R \quad . \tag{3.4}$$

Hence, by using Lemma 3.1.5, we can define the operator $T(1)$ on $\mathcal{D}_0(\mathbb{R}^n)$ by the following relation,

$$\langle T(1), f\rangle = \langle T^t f, 1\rangle \quad .$$

$T(1)$ is thus a continuous linear form on the subspace $\mathcal{D}_0(\mathbb{R}^n) \subset \mathcal{D}(\mathbb{R}^n)$. We extent $T(1)$ to a distribution in $\mathcal{D}'(\mathbb{R}^n)$ as follows.

Let $\phi \in \mathcal{D}(\mathbb{R}^n)$ be a fixed function of mean 1: then every function $f \in \mathcal{D}(\mathbb{R}^n)$ can be written as $f = \lambda\phi + g$, where $\lambda = \int f(x)dx$ and $g \in \mathcal{D}_0(\mathbb{R}^n)$. We define a new operator S on $\mathcal{D}(\mathbb{R}^n)$ by $\langle S, f\rangle := \lambda\alpha + \langle T(1), g\rangle$, where $\alpha \in \mathbb{R}$ is fixed. Then we extend $T(1)$ to a distribution in $\mathcal{D}'(\mathbb{R}^n)$.

Another possible definition of $T(1)$ is given by the following direct approach. We start with a function $\phi \in \mathcal{D}(\mathbb{R}^n)$ which equals 1 at 0 and put $\phi_\epsilon(x) = \phi(\epsilon x)$, for every $\epsilon > 0$. We then consider the distribution $S_\epsilon = T(\phi_\epsilon)$.

Lemma 3.1.6 *There exists constants* $c(\epsilon)$ *such that* $\lim_{\epsilon\downarrow 0}(S_\epsilon - c(\epsilon))$ *exists in the sense of distributions and the limit is defined to be* $T(1)$.

Indeed, if $f \in \mathcal{D}_0(\mathbb{R}^n)$, we have $\langle T(\phi_\epsilon), f\rangle = \langle \phi_\epsilon, T^t f\rangle$. We must verify that the right-hand side converges to $\langle 1, T^t f\rangle$. Putting $1 = u + v$, where $u \in \mathcal{D}(\mathbb{R}^n)$ and $u = 1$ in a compact neighborhood of the support of f. Since the support of u is compact, we can handle $\lim_{\epsilon\downarrow 0}\langle \phi_\epsilon, uT^t f\rangle$, using the definition of a distribution. Moreover, $T^t f$ is integrable by (3.4), the Lebesgue dominated convergence theorem then shows that $\lim_{\epsilon\downarrow 0}\langle \phi_\epsilon, vT^t f\rangle$ exists.

Let $w \in \mathcal{D}(\mathbb{R}^n)$ be a function of mean 1 and put $c(\epsilon) = \langle S_\epsilon, w\rangle$. Then for $f \in \mathcal{D}(\mathbb{R}^n)$ and $\lambda = \int f(x)dx$,

$$\langle S_\epsilon - c(\epsilon), f\rangle = \langle S_\epsilon, f\rangle - \lambda\langle S_\epsilon, w\rangle = \langle S_\epsilon, g\rangle \quad ,$$

where $g = f - \lambda w$. Then $g \in \mathcal{D}_0(\mathbb{R}^n)$, so the required limit exists.

3.2 The Wavelet Proof of the $T(1)$-Theorem

We need the definition of BMO to state the $T(1)$-Theorem.

Definition 3.2.1 *Let* f *be a locally integrable function and* Q *be cube on* $\mathbb{R}^n$. *Let*

$$M_Q f = \frac{1}{|Q|}\int_Q f(x)dx \quad . \tag{3.5}$$

A locally integrable function f is said to be bounded mean oscillation $\left(BMO(\mathbb{R}^n)\right)$ if there is a constant $A > 0$ such that

$$\frac{1}{|Q|}\int_Q |f(x) - M_Q f|dx \le A \tag{3.6}$$

holds for all cube Q. The smallest A for which (3.6) is satisfied is then taken to be the norm of f in this space, and is denoted by $\|f\|_{BMO}$.

The notion of bounded mean oscillation was first introduced in 1961 by F. John and L. Nirenberg [19]. The important development of BMO started when Fefferman and Stein discovered the fact that BMO is the dual space of the Hardy space H^1 ([11], [12]).

Theorem 3.2.2 *Let $T : V \to V'$ be the singular integral as in Definition 3.1.3. Then a necessary and sufficient condition for the extension of T as a continuous linear operator on $L^2(\mathbb{R}^n)$ is that all the following properties are satisfied:*

(i) $T(1)$ belongs to $BMO(\mathbb{R}^n)$;

(ii) $T^t(1)$ belongs to $BMO(\mathbb{R}^n)$;

(iii) T is weakly continuous on $L^2(\mathbb{R}^n)$.

It was proved that in [30], [33], [36] that Calderón-Zygmund operators admit bounded extensions from L^∞ to BMO . The function $1 \in L^\infty$, therefore (i) is satisfied. The transpose (or adjoint) of a Calderón-Zygmund operator is still a Calderón-Zygmund operator. So (ii) is satisfied. Finally, any operator which is L^2 continuous is automatically weakly continuous. Thus, we need only to prove the sufficiency of the $T(1)$-Theorem.

The $T(1)$-Theorem becomes stronger as the exponent $\gamma > 0$ in the definition of $K(x,y)$ decrease. We may therefore suppose that $0 < \gamma < 1$ in what follows.

The wavelets ψ_λ are real-valued and the idea of the proof is to estimate the absolute value $|\tau(\lambda,\lambda')|$ of the entries $\tau(\lambda,\lambda') = \langle T\psi_\lambda, \psi_{\lambda'}\rangle = \langle \psi_\lambda, T^t\psi_{\lambda'}\rangle$, which is the matrix representation of T with respect to the wavelet basis $\{\psi_\lambda\}_{\lambda\in\Lambda}$. And we want to show that these entries become very small when the cube $Q(\lambda)$ and $Q(\lambda')$ differ either in position or magnitude. We first suppose $T(1) = T^t(1) = 0$ to see the rate of decay of $|\tau(\lambda,\lambda')|$.

Lemma 3.2.3 *Let $f(x)$ be a C^1 function whose support is contained in the unit ball of $\mathbb{R}^n$ and whose partial derivatives $\partial f/\partial x_j$, $1 \le j \le n$, satisfy $\|\partial f/\partial x_j\|_\infty \le 1$. Let $\gamma \in (0,1)$ be an exponent and let $g \in L^1(\mathbb{R}^n)$ be a function satisfying $|g(x)| \le (1+|x|)^{-n-\gamma}$ and $\int g(x)dx = 0$. Then there is a constant $C = C(n,\gamma)$ such that, for all $x_0 \in \mathbb{R}^n$ and every $r \ge 1$,*

$$\left|\int g(x) f\Big(\frac{x-x_0}{r}\Big)dx\right| \le Cr^{-\gamma} \ , \quad \text{if } |x_0| \le r \ , \tag{3.7}$$

and

$$\left|\int g(x) f\Big(\frac{x-x_0}{r}\Big)dx\right| \le Cr^{n}|x_0|^{-n-\gamma} \ , \quad \text{if } |x_0| \ge r \ . \tag{3.8}$$

To establish (3.7) and (3.8), since $g \in L^1(\mathbb{R}^n)$, we can apply the Lebesgue differentiation theorem to write

$$g(x) = \frac{\partial}{\partial x_1} g_1(x) + \cdots + \frac{\partial}{\partial x_n} g_n(x) \ ,$$

where $|g_1(x)| \le C(1+|x|)^{-n-\gamma+1}, \cdots, |g_n(x)| \le C(1+|x|)^{-n-\gamma+1}$ and then integrate by parts.

Proposition 3.2.4 *Let $T: V \to V'$ be a weakly continuous operator on $L^2(\mathbb{R}^n)$ which corresponds to a singular integral, as in Definition 3.1.3. Suppose, further, that $T(1) = T^t(1) = 0$. Let $\{\psi_\lambda\}_{\lambda\in\Lambda}$ be an orthonormal C^1 wavelet basis of compact support. Then there exists a constant C such that, for $\lambda = 2^{-j}k + 2^{-j-1}\epsilon$ and $\lambda' = 2^{-j'}k' + 2^{-j'-1}\epsilon'$,*

$$|\tau(\lambda,\lambda')| \le C 2^{-|j-j'|((n/2)+\gamma)} \left(\frac{2^{-j}+2^{-j'}}{2^{-j}+2^{-j'}+|2^{-j}k - 2^{-j'}k'|}\right)^{n+\gamma} . \tag{3.9}$$

Proof: Firstly, we observe that everything is symmetric in λ and λ'. We may therefore suppose that $j \leq j'$.

If $j = j'$, when the support of ψ_λ and $\psi_{\lambda'}$ intersect, (3.9) follows from the weak continuity of T. If the supports are disjoint, we have

$$\begin{aligned}\tau(\lambda, \lambda') &= \int\int K(x,y)\psi_{\lambda'}(x)\psi_\lambda(y)\, dxdy \\ &= \int\int \Big(K(x,y) - K(x,\lambda)\Big)\psi_{\lambda'}(x)\psi_\lambda(y)\, dxdy \ ,\end{aligned}$$

and we get the upper bound from (iii) of Definition 3.1.3.

If $j' \geq j+1$, the wavelet $\psi_{\lambda'}$ are in V_j. Before computing $\langle T\psi_\lambda, \psi_{\lambda'}\rangle$, we use the orthogonal projection of $T\psi_\lambda$ on V_j. Note that this would not change the scalar product. Since $\{\phi_{jl}\}_{l\ \in\mathbb{Z}^n}$ is an orthonormal basis of the subspace V_j of $L^2(\mathbb{R}^n)$, we replace the distribution $T\psi_\lambda$ by

$$g(x) = \sum_{l\in\mathbb{Z}^n} c(l) 2^{nj/2}\phi(2^j x - l) \ , \tag{3.10}$$

where $c(l) = \langle T\psi_\lambda, \phi_{jl}\rangle$ and $\phi_{jl}(x) = 2^{nj/2}\phi(2^j x - l)$.

The coefficients $c(l)$ can be dealt with by using the weak continuity of T and the estimate (iii) of the kernel $K(x,y)$ in Definition 3.1.3. In this way, we get

$$|c(l)| = |\langle T\psi_\lambda, \phi_{jl}\rangle| \leq C(1 + |k-l|)^{-n-\gamma} \ . \tag{3.11}$$

Further, since $\sum \phi(x-l) = 1$ and $T^t(1) = 0$, we get $\sum_{l\in\mathbb{Z}^n} c(l) = 0$.

Note that $g(x)$ defined by equation (3.10) satisfies $g \in L^1(\mathbb{R}^n)$, $\int g(x)dx = 0$. Moreover, by (3.11) and that $\phi(x)$ has compact support, we obtain $|g(x)| \leq C(1+|x|)^{-n-\gamma}$. Hence $g(x)$ in (3.10) satisfies all the assumptions in Lemma 3.2.3.

Moreover, the sup-norm of all the partial derivatives of $\psi_{\lambda'}$ are small. All this facts allow us to make use of Lemma 3.2.3. The estimate (3.9) of the case $j' \geq j+1$ follows immediately. $\square$

The following is the Schur's lemma [6].

Lemma 3.2.5 *Let $M = (m(p,q))_{p,q\in\mathbb{N}}$ be an infinite matrix and suppose that $w(p) > 0$ is a sequence of positive real numbers. Suppose, further, that, for every $p \in \mathbb{N}$,*

$$\sum_{q\in\mathbb{N}} |m(p,q)|w(q) \leq \beta w(p) \tag{3.12}$$

and, symmetrically, for every $q \in \mathbb{N}$,

$$\sum_{p\in\mathbb{N}} |m(p,q)|w(p) \leq \alpha w(q). \tag{3.13}$$

Then $M : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ is bounded and $\|M\| \leq \beta\alpha$.

Now we can prove the $T(1)$-Theorem under an additional hypothesis that $T(1) = T^t(1) = 0$.

Proposition 3.2.6 *Let $T : V \to V'$ be as in Theorem 3.2.2 be a weakly continuous operator on $L^2(\mathbb{R}^n)$ and if $T(1) = T^t(1) = 0$, then T is a bounded linear operator on $L^2(\mathbb{R}^n)$.*

Proof: We apply Schur's lemma with $w(\lambda) = 2^{-nj/2}$ and using (3.9) to estimate the following sum in Proposition 3.2.4,

$$\sum_{j'}\sum_{k'} 2^{-nj'/2}2^{-|j-j'|((n/2)+\gamma)}\left(\frac{2^{-j}+2^{-j'}}{2^{-j}+2^{-j'}+|2^{-j}k-2^{-j'}k'|}\right)^{n+\gamma} \tag{3.14}$$

We start with the case $j' \geq j$. Putting $d = j' - j$, we see that the quotient above is bounded by $2(1+|k-2^{-d}k'|)^{-1}$. Hence,

$$\sum_{k'} 2^{-nd}(1+|k-2^{-d}k'|)^{-n-\gamma} = \sum_{k'} 2^{-nd}(1+2^{-d}|k'|)^{-n-\gamma} \leq C(n,\gamma),$$

as the latter series can be interpreted as a Riemann sum. This leaves

$$\left(\sum_{j'\geq j} 2^{-\gamma(j'-j)}\right) 2^{-nj/2} = C(\gamma)2^{-nj/2} = C(\gamma)w(\lambda) \tag{3.15}$$

Similarly, if $j' < j$, we put $d = j - j'$ and the quotient is bounded above by $2(1+|2^{-d}k - k'|)^{-1}$. This leads to a summation over k' which is bounded above, uniformly with respect to k and d. We are left with

$$\sum_{j'<j} 2^{-nj'/2} 2^{-(j-j')(n/2+\gamma)} = 2^{-nj/2} \sum_{d>0} 2^{-d\gamma} = C(\gamma) 2^{-nj/2}. \tag{3.16}$$

Since everything is symmetric with respect to j and k, we have proved that (3.14) is finite and the matrix $\{\tau(\lambda, \lambda')\}_{\lambda,\lambda'\in\Lambda}$ is bounded on $l^2(\Lambda)$ and the proof is completed. □

3.3 Proof of the $T(1)$-Theorem (Continue)

We have just proved the $T(1)$-theorem under the hypothesis that $T(1) = T^t(1) = 0$, i.e., the integrals

$$\alpha(\lambda) = \int T^t(\psi_\lambda)\, dx = \langle T(1), \psi_\lambda \rangle = 0 \ , \tag{3.17}$$

$$\beta(\lambda) = \int T(\psi_\lambda)\, dx = \langle T^t(1), \psi_\lambda \rangle = 0 \ , \tag{3.18}$$

for all $\lambda \in \Lambda$. In order to eliminate this assumption. Let us recall a result in Y. Meyer [26].

Proposition 3.3.1 *Let $b(x)$ be a function belonging to BMO. Then its wavelet coefficient $a(\lambda) = \langle b, \psi_\lambda \rangle$ satisfy the Carleson's condition :*

For $\lambda = 2^{-j}k + 2^{-j-1}\epsilon$, let $Q(\lambda)$ denote the cube defined by $2^j x - k \in [0,1)^n$ as before. Then there exists a constant C such that, for each dyadic cube Q in $\mathbb{R}^n$,

$$\sum_{Q(\lambda)\subset Q} |a(\lambda)|^2 \le C|Q| \ . \tag{3.19}$$

Proof: Let $b(x) \in BMO$. The support of the wavelet ψ_λ is contained in a cube Γ which has the same center as the cube $Q(\lambda)$ but is m times as big (m is independent of λ). Write $b(x)$ as

$$b(x) = b_1(x) + b_2(x) + M_\Gamma b \ ,$$

where $M_\Gamma b$ is the mean of $b(x)$ on Γ as in (3.5) and

$$b_1(x) = \begin{cases} b(x) - M_\Gamma b & \text{if } x \in \Gamma , \\ 0 & \text{otherwise} . \end{cases}$$

Note also that if $Q(\lambda) \subset Q$, the coefficients $\langle b_2, \psi_\lambda \rangle = 0$. On the other hand, the integral of each wavelet ψ_λ is zero. So $\langle b, \psi_\lambda \rangle = \langle b_1, \psi_\lambda \rangle$, we have

$$\sum_{Q(\lambda) \subset Q} |\langle b, \psi_\lambda \rangle|^2 \leq \sum_{\lambda \in \Lambda} |\langle b_1, \psi_\lambda \rangle|^2 = \|b_1\|_{L^2}^2 \leq \left(m^n \|b\|_{BMO}^2 \right) |Q| \, . \quad \square$$

By the proposition above, we have the following inequalities for BMO: for every dyadic cube Q,

$$\sum_{Q(\lambda) \subset Q} |\alpha(\lambda)|^2 \leq C|Q| \ , \tag{3.20}$$

$$\sum_{Q(\lambda) \subset Q} |\beta(\lambda)|^2 \leq C|Q| \ . \tag{3.21}$$

The inequalities (3.20) and (3.21) will enable us to construct two auxiliary Calderón-Zygmund operators R and S such that

$$R(1) = T(1), \quad R^t(1) = 0, \quad S(1) = 0, \quad S^t(1) = T^t(1). \tag{3.22}$$

The details will be given in the rest of the section. By using this, we put $N = T - R - S$ and then

(i) $N : V \to V'$ is weakly continuous and its kernel satisfies the conditions in Definition 3.1.3, because all three operators R, S and T have these properties.

(ii) $N(1) = N^t(1) = 0$.

N is a bounded linear operator on $L^2(\mathbb{R}^n)$. Since the same is true for R and S. It follows that T is a bounded linear operator on $L^2(\mathbb{R}^n)$ by Proposition 3.2.6 and the proof of the $T(1)$-Theorem is complete.

To construct the operators R and S, we first observe that

Lemma 3.3.2 *Let $\theta \in \mathcal{D}(\mathbb{R}^n)$ be a function of mean 1 and put $\theta_\lambda(x) = 2^{nj}\theta(2^j x - k)$, where $\lambda = k2^{-j} + \epsilon 2^{-j-1}$ and $\epsilon \in \{0,1\}^n \backslash \{(0,\cdots,0)\}$ as before. Let $\{u(\lambda)\}_{\lambda\in\Lambda}$ be a sequence of complex number such that $|u(\lambda)| \leq C2^{-nj/2}$, for some constant C. Then*

$$H(x,y) := \sum_{\lambda\in\Lambda} u(\lambda)\psi_\lambda(x)\theta_\lambda(y) \tag{3.23}$$

is a distribution, i.e. $H(x,y) \in \mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n)$.

Proof: To see that $H(x,y)$ is a distribution, it suffices to show that $H \in L^1_{\text{loc}}(\mathbb{R}^n \times \mathbb{R}^n)$, i.e.,

$$\int_{Q_N}\int_{Q_N} |H(x,y)|\, dxdy < \infty \quad , \ \forall N \in \mathbb{N} \ , \tag{3.24}$$

where Q_N is the cube in $\mathbb{R}^n$ center at the origin of sides length $2N$. Note that

$$\begin{aligned} &\int_{Q_N}\int_{Q_N} |H(x,y)|\, dxdy \\ \leq\ & \sum_{\lambda\in\Lambda} |u(\lambda)| \int_{Q_N} |\psi_\lambda(x)|dx \int_{Q_N} |\theta_\lambda(y)|dy \\ =\ & \sum_{\epsilon}\sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^n} |u(\lambda)| \int_{Q_N} |\psi_\lambda(x)|dx \int_{Q_N} |\theta_\lambda(y)|dy \ . \end{aligned} \tag{3.25}$$

Let us decompose (3.25) into two parts as

$$\begin{aligned} &\sum_{j=0}^{\infty}\sum_{k\in\mathbb{Z}^n} |u(\lambda)| \int_{Q_N} |\psi_\lambda(x)|dx \int_{Q_N} |\theta_\lambda(y)|dy \\ +\ & \sum_{j=-1}^{-\infty}\sum_{k\in\mathbb{Z}^n} |u(\lambda)| \int_{Q_N} |\psi_\lambda(x)|dx \int_{Q_N} |\theta_\lambda(y)|dy \ . \end{aligned} \tag{3.26}$$

The decaying property of the sequence $\{u(\lambda)\}$ when $j \geq 0$in the hypothesis implies the first sum of (3.26) is finite. For the second sum, note that when j decreases, the measure of the support of ψ_λ will increase. Recall that the support of ψ_λ is contained in a cube Γ as in the proof of Proposition 3.3.1 of volume $m2^{-jn}$. So, there exists largest $j_0 < 0$ such that for $j < j_0$ there is one and only one ψ_λ with support intersecting Q_N. This means that for each $j < j_0$, it corresponds to one and only one $k \in \mathbb{Z}^n$ such that the integral $\int_{Q_N} |\psi_\lambda(x)|dx$ is non zero.

On the other hand, as j decreases, the sup-norm of ψ_λ will decrease as well. Let's compute the second sum of (3.26). It's obvious that

$$\sum_{j=-1}^{j_0} \sum_{k \in \mathbb{Z}^n} |u(\lambda)| \int_{Q_N} |\psi_\lambda(x)|dx \int_{Q_N} |\theta_\lambda(y)|dy$$

is finite, since for each j such that $j_0 \leq j \leq -1$, it corresponds to finitely many $k \in \mathbb{Z}^n$ such that the support of ψ_λ intersect Q_N. Moreover

$$\begin{aligned}
& \sum_{j=j_0-1}^{-\infty} \sum_{k \in \mathbb{Z}^n} |u(\lambda)| \int_{Q_N} |\psi_\lambda(x)|dx \int_{Q_M} |\theta_\lambda(y)|dy \\
\leq\ & C \sum_{j=j_0-1}^{-\infty} 2^{-nj/2} \cdot 2^{nj/2} \int_{Q_N} |\psi_\epsilon(2^j x - k)|dx \cdot 2^{nj} \int_{Q_N} |\theta(2^j y - k)|dy \\
\leq\ & C \sum_{j=j_0-1}^{-\infty} 2^{nj} \cdot 2|Q_N| \cdot \|\psi_\epsilon\|_\infty \cdot \|\theta\|_\infty \\
<\ & \infty \ ,
\end{aligned}$$

where $|Q_N|$ denote the volume of Q_N. This complete our proof that $H(x,y)$ defined in (3.23) belongs to $\mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n)$. $\square$

Remark: By putting $Q = Q(\lambda)$ into (3.20) and (3.21), we have $|\alpha(\lambda)| \leq C2^{-nj/2}$ and $\beta(\lambda) \leq C2^{-nj/2}$, where $\alpha(\lambda)$ and $\beta(\lambda)$ are defined in (3.17) and (3.18) respectively. Hence, we can apply the Lemma 3.3.2 to the sequences $\{\alpha(\lambda)\}_{\lambda \in \Lambda}$ and $\{\beta(\lambda)\}_{\lambda \in \Lambda}$.

The following lemma is due to Carleson [37].

Lemma 3.3.3 *Let $p(\lambda)$, $\lambda \in \Lambda$ be a sequence of positive numbers such that*

$$\sum_{Q(\lambda) \subset Q} p(\lambda) \leq |Q|$$

for every dyadic cube Q. Then for every sequence $w(\lambda) \geq 0$, $\lambda \in \Lambda$, we have

$$\sum_{\lambda \in \Lambda} w(\lambda) p(\lambda) \leq \int_{\mathbb{R}^n} \widetilde{w}(x) dx, \tag{3.27}$$

where $\widetilde{w}(x) = \sup_{Q(\lambda)\ni x} w(\lambda)$.

Proof: Let

$$\chi(\lambda,t) = \begin{cases} 1 & \text{if } 0 < t < w(\lambda) \\ 0 & \text{otherwise} \end{cases}$$

Then $\sum_{\lambda\in\Lambda} w(\lambda)p(\lambda) = \int_0^\infty \sum_{\lambda\in\Lambda} \chi(\lambda,t)p(\lambda)dt$. For $t > 0$, let

$$\Omega_t = \{x : \widetilde{w}(x) > t\} = \cup\{Q(\lambda) : w(\lambda) > t\} .$$

Using Chebyshev's inequality [20] and observed that $\int_0^\infty |\Omega_t|dt = \int_{\mathbb{R}^n} \widetilde{w}(x)dx$, we have

$$|\Omega_t| \le t^{-1} \int_{\mathbb{R}^n} \widetilde{w}(x)dx,$$

We may suppose that this integral is finite, because, if not, (3.27) is trivially true.

Let Q_k denote the maximal dyadic cubes contained in Ω_t. Ω_t is the union of the Q_k and, for $t > 0$, we have

$$\sum_{\lambda\in\Lambda} \chi(\lambda,t)p(\lambda) \le \sum_{Q(\lambda)\subset\Omega_t} p(\lambda) = \sum_k \sum_{Q(\lambda)\subset Q_k} p(\lambda) \le \sum_k |Q_k| = |\Omega_t| .$$

Thus

$$\begin{aligned} \sum_{\lambda\in\Lambda} w(\lambda)p(\lambda) &= \int_0^\infty \sum_{\lambda\in\Lambda} \chi(\lambda,t)p(\lambda)dt \\ &\le \int_0^\infty |\Omega_t|dt \\ &= \int_{\mathbb{R}^n} w(x)dx. \quad \square \end{aligned}$$

By using Carleson's lemma and Proposition 3.3.1, we obtain the following proposition which yields the construction of the operators R and S.

Proposition 3.3.4 *Let* $T : \mathcal{D}(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$ *be an integral operator with kernel* $H(x,y)$ *as in (3.23) and is weakly continuous on* $L^2(\mathbb{R}^n)$. *Then* T *can be extended as a continuous linear operator on* $L^2(\mathbb{R}^n)$ *if and only if* $\sum_{\lambda\in\Lambda} \alpha(\lambda)\psi_\lambda(x) \in$

BMO, or, equivalently, if there is a constant C such that, for every dyadic cube Q,

$$\sum_{Q(\lambda)\subset Q} |\alpha(\lambda)|^2 \le C|Q| \ . \tag{3.28}$$

Proof: Since T satisfies (i), (ii), (iii) in the Definition 3.1.3, $T(1) \in BMO$ and $\int \theta_\lambda(y)dy = 1$. By Proposition 3.3.1, we obtain the following inequality

$$\sum_{Q(\lambda)\in Q} |\alpha(\lambda)|^2 \le C|Q| \ ,$$

which show that (3.28) is necessary.

Let $f \in L^2(\mathbb{R}^n)$. Then

$$T(f) = \sum_{\lambda\in\Lambda} \alpha(\lambda)\theta_\lambda(f)\psi_{la}(x),$$

where $\theta_\lambda(f) = \int f(y)\theta_\lambda(y)dy$. Since the wavelets ψ_λ form an orthonormal basis for $L^2(\mathbb{R}^n)$, we have

$$\|T(f)\|_2^2 = \sum_{\lambda\in\Lambda} |\alpha(\lambda)|^2|\theta_\lambda(f)|^2.$$

Now, let $p(\lambda) = |\alpha(\lambda)|^2$, $w(\lambda) = |\theta_\lambda(f)|^2$ and apply Carleson's lemma, we obtain

$$\sum_{\lambda\in\Lambda} |\alpha(\lambda)|^2|\theta_\lambda(f)|^2 \le \int_{\mathbb{R}^n} w(x)dx,$$

where $w(x) = f^*(x)^2$ and $f^*(x)$ is the Hardy-Littlewood maximal function of f. Since $\int_{\mathbb{R}^n} f^*(x)^2dx \le C\|f\|_2^2$. We finally get

$$\|T(f)\|_2^2 \le C\|f\|_2^2,$$

which show that (3.28) is sufficient. □

Let's summarize our founding in the following proposition.

Proposition 3.3.5 *Let $\alpha(\lambda)$ and $\beta(\lambda)$ be defined as in (3.17) and (3.18). Let R and S be integral operators defined by kernels $H_1(x,y) = \sum_{\lambda\in\Lambda} \alpha(\lambda)\psi_\lambda(x)\theta_\lambda(y)$*

and $H_2(x,y) = \sum_{\lambda\in\Lambda}\beta(\lambda)\theta_\lambda(x)\psi_\lambda(y)$ *respectively, as in Lemma 3.3.2. Then* R *and* S *are the Calderón-Zygmund operators satisfies* $R(1) = T(1)$,$R^t(1) = 0$, $S(1) = 0$ *and* $S^t(1) = T^t(1)$.

The above proposition give us the operators R and S we have been try to construct. The proof of the $T(1)$-Theorem is complete by previous remarks on the operator N.

3.4 Some recent results on the $T(1)$-Theorem

The kernel conditions in Definition 3.1.3 was first weakened by Y. Meyer [28], who replace the kernel estimates in the Definition 3.1.3 (i), (ii), (iii) by the integral estimates:

$$\sup_{r\geq 0}\int_{r\leq|x-y|\leq 2r}(|K(x,y)|+|K(y,x)|)\,dy < C, \tag{3.29}$$

and

$$\sup_{r>0,|u|+|v|\leq r}\{\int_{2^kr\leq|x-y|\leq 2^{k+1}r}(|K(x+u,y+v)-K(x,y)|)dx$$
$$\int_{2^kr\leq|x-y|\leq 2^{k+1}r}(|K(x+u,y+v)-K(x,y)|)dy\} = B(k), \tag{3.30}$$

where $k = 1,2,3,\cdots$, and $B(k)$ satisfy

$$\sum_{k=1}^{\infty} kB(k) < \infty. \tag{3.31}$$

In [16], Y. S. Han and S. Hofmann consider a slightly weaker version of Meyer's conditions (3.30):

$$\sup_{r>0,|u|+|v|\leq r}\{\int_{2^jr\leq|x-y|}(|K(x+u,y+v)-K(x,y)|)dx$$
$$+\int_{2^jr\leq|x-y|}(|K(x+u,y+v)-K(x,y)|)dy\} = \gamma(j), \tag{3.32}$$

where

$$\sum_{j=1}^{\infty}\gamma(j) < \infty. \tag{3.33}$$

Since $\gamma(j) \leq \sum_{k=j}^{\infty} B(k)$, it is easy to see that (3.31) implies (3.33). Of course, if $B(j)$ or $\gamma(j) \leq C2^{-j\epsilon}$ for some $\epsilon > 0$, then (3.31) and (3.33) are equivalent.

In [10], Deng Donggao, Yan Lixin and Yang Qixiang obtain a stronger result which improve both Meyer's and Han-Hofmann's conditions in some sense. The main result of Deng , Yan and Yang in [10] is as follows:

Theorem 3.4.1 *Let T be an operator as in Definition 3.1.3 (iv), and it associated with a kernel $K(x,y)$ satisfying (3.29) and*

$$\sum_{k=1}^{\infty} \sqrt{k} B(k) < \infty \tag{3.34}$$

where $B(k)$ is defined as in (3.30). Then T can be extend to a bounded operator on L^2 if and only if $T(1) \in BMO$, $T^t(1) \in BMO$ and T satisfies

$$|\langle T\phi, \psi\rangle| \leq Ct^n[\|\phi\|_\infty + t\|\nabla\phi\|_\infty][\|\psi\|_\infty + t\|\nabla\psi\|_\infty], \tag{3.35}$$

for and $\phi, \psi \in \mathcal{D}$ and diam $\mathrm{supp}\phi \leq t$, *diam* $\mathrm{supp}\psi \leq t$.

Chapter 4

Singular Values of Compact Pseudodifferential Operators

This chapter discuss two results of Heil, Ramanathan and Topiwala [17]. The first one is about the asymptotic decay of the singular values of compact operators arising from the Weyl correspondence. The Weyl correspondence is a formalism that bijectively associates to any continuous linear operator $L : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ a distributional symbol $\sigma \in \mathcal{S}'(\mathbb{R}^n)$. The motivating problem is to find sufficient conditions on symbol which ensure that the corresponding operator has singular values with a prescribed rate of decay.

The second one is a new development and improvement of the Calderón-Vaillancourt Theorem in the context of the Weyl Correspondence. The classical Calderón-Vaillancourt Theorem for the Weyl correspondence states that the pseudodifferential operator L_σ is a bounded operator on $L^2(\mathbb{R}^n)$ if $\sigma \in C^{2n+1}(\mathbb{R}^{2n})$. In the following, we will discuss the theorem in a more general form, which stated in terms of Hölder-Zygmund classes $\Lambda^s(\mathbb{R}^n)$ such that L_σ is bounded on $L^2(\mathbb{R}^n)$ if $\sigma \in \Lambda^{2n+\epsilon}(\mathbb{R}^{2n})$ for all $\epsilon > 0$.

The Weyl correspondence plays an important role in a variety of contexts, including quantum mechanics and partial differential equations. A detail description of the Weyl calculus can be found in [13].

4.1 Background

4.1.1 Singular Values

Let X and Y be normed spaces. An linear operator $T : X \to Y$ is called compact if for every bounded subset M of X, the image $T(M)$ is relatively compact, that is the closure $\overline{T(M)}$ is compact.

Let H be a separable Hilbert space and $\mathcal{B}(H)$ be the space of all bounded linear operators mapping H into itself equipped with the usual operator norm $\|\cdot\|_{\mathcal{B}(H)}$. Let L^* be the adjoint of L, then L^*L is a positive operator. Let $L \in \mathcal{B}(H)$ be a compact operator, then L^*L is compact as well and has a discrete spectrum tending towards zero. We define the *singular value* $s_k(L)$ of L to be the square root of the k-th largest eigenvalue of L^*L. Let $\operatorname{rank}(T)$ denote the dimension of the range of the operator T and let

$$a_k(L) = \inf\{\|L - T\|_{\mathcal{B}(H)} : \operatorname{rank}(T) < k\} \tag{4.1}$$

be the *approximation numbers* of L. Since H is a separable Hilbert space, the singular values of a compact operator L coincide with the approximation numbers of L, i.e., $s_k(L) = a_k(L)$ (see [15] for the proof).

4.1.2 Schatten Class $\mathcal{I}_p$

One way of quantifying the rate of decay of the singular values of a compact operator L is by determining the l^p class to which they belong. This leads to the definition of the *Schatten class* $\mathcal{I}_p$ as the set of all compact operators $L \in \mathcal{B}(H)$ for which the sequence of singular values $\{s_k(L)\}$ is in l^p. In particular, $\mathcal{I}_\infty$ is the space of all compact operators on H. Other useful identifications are that $\mathcal{I}_1$ is the space of all trace-class operators on H, and $\mathcal{I}_2$ is the space of all Hilbert-Schmidt operators on H.

The Schatten class $\mathcal{I}_p$ is a Banach space under the norm

$$\|L\|_{\mathcal{I}_p} = \|\{s_k(L)\}\|_{l^p} = \begin{cases} \left(\sum_k s_k(L)^p\right)^{1/p} & 1 \leq p < \infty \ , \\ \sup_k s_k(L) = s_1(L) = \|L\|_{\mathcal{B}(H)} & p = \infty \ . \end{cases} \tag{4.2}$$

The singular values of two compact operators L_1, L_2 obey the inequality

$$s_{k+l+1}(L_1 + L_2) \leq s_{k+1}(L_1) + s_{l+1}(L_2) \ ,$$

As a consequence, the following refinement of (4.1) holds for $L \in \mathcal{I}_2$:

$$\sum_{k>N} s_k(L)^2 \leq \inf\{\|L - T\|_{\mathcal{I}_2}^2 : \operatorname{rank}(T) \leq N\} \ . \tag{4.3}$$

This inequality will play a key role in our later estimates.

4.1.3 The Ambiguity Function and the Wigner Distribution

Let $\mathcal{S}(\mathbb{R}^n)$ be the set of all infinitely differentiable functions ϕ on $\mathbb{R}^n$ such that all their derivatives remain bounded when multiplied by arbitrary polynomials. One can define a denumerable collection of seminorms $\|\cdot\|_{\alpha,\beta}$ on $\mathcal{S}(\mathbb{R}^n)$ as

$$\|\phi\|_{\alpha,\beta} = \sup_{x\in\mathbb{R}^n} |x^\alpha D_x^\beta \phi(x)| \ .$$

Here we use conventional notation:

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \ , \qquad D_x^\beta = \frac{\partial^{\beta_1}}{\partial x_1^{\beta_1}} \cdots \frac{\partial^{\beta_n}}{\partial x_n^{\beta_n}} \ , \tag{4.4}$$

$\alpha = (\alpha_1, \cdots, \alpha_n)$ and $\beta = (\beta_1, \cdots, \beta_n)$ are n-tuples of natural numbers. The function space $\mathcal{S}(\mathbb{R}^n)$ is called the *Schwartz class* of functions.

A *tempered distribution* is a linear functional on $\mathcal{S}(\mathbb{R}^n)$ that is continuous in the topology on $\mathcal{S}(\mathbb{R}^n)$ induced by this family of seminorms. The set of all tempered distributions with this weak topology is denoted by $\mathcal{S}'(\mathbb{R}^n)$

In order to define the Weyl correspondence, we need the *(cross-)ambiguity function* and the *(cross-)Wigner distribution* of $f, g \in L^2(\mathbb{R}^n)$.

Definition 4.1.1 *Let $\alpha = (p, q) \in \mathbb{R}^n \times \mathbb{R}^n$ and let*

$$\rho(\alpha) f(x) = \rho(p, q) f(x) = e^{\pi i p q} e^{2\pi i q x} f(x + p) \tag{4.5}$$

be the time-frequency shift of f. The (cross-)ambiguity function A of $f, g \in L^2(\mathbb{R}^n)$ is defined by the following relation:

$$\begin{aligned} A(f,g)(p,q) &= \langle \rho(p,q) f, g \rangle \\ &= e^{\pi i p q} \int e^{2\pi i q x} f(x+p) \overline{g(x)} \, dx \\ &= \int e^{2\pi i q x} f(x + \frac{p}{2}) \overline{g(x - \frac{p}{2})} \, dx \ . \end{aligned} \tag{4.6}$$

It is clear from the Schwarz inequality that $A(f, g)$ is always a bounded, continuous function on $\mathbb{R}^{2n}$ satisfying $\|A(f,g)\|_\infty \leq \|f\|_{L^2(\mathbb{R}^n)} \|g\|_{L^2(\mathbb{R}^n)}$.

The function A can be extended in an obvious way from a sesquilinear map defined on $L^2(\mathbb{R}^n) \times L^2(\mathbb{R}^n)$ to a linear map $\widetilde{A}$ defined on the tensor product $L^2(\mathbb{R}^n) \otimes L^2(\mathbb{R}^n)$, which is naturally isomorphic to $L^2(\mathbb{R}^{2n})$. Namely, if $F \in L^2(\mathbb{R}^{2n})$, we define

$$\widetilde{A}(F)(p,q) = \int e^{2\pi i q y} F(x + \frac{1}{2}p, x - \frac{1}{2}p) \, dx \ . \tag{4.7}$$

We then have $A(f, g) = \widetilde{A}(f \otimes \overline{g})$, where $f \otimes \overline{g}(x, y) = f(x)\overline{g(y)}$. $\widetilde{A}$ is the composition of the measure-preserving change of variables $(x,p) \to (x + \frac{1}{2}p, x - \frac{1}{2}p)$ with inverse Fourier transformation in the first variable. Therefore it is unitary on $L^2(\mathbb{R}^{2n})$, maps $\mathcal{S}(\mathbb{R}^{2n})$ onto itself, and extends to a continuous bijection $\mathcal{S}'(\mathbb{R}^{2n})$ onto itself. Analogous of these facts transfer back to A on $L^2(\mathbb{R}^n) \times L^2(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^n)$ and $\mathcal{S}'(\mathbb{R}^n) \times \mathcal{S}'(\mathbb{R}^n)$.

Definition 4.1.2 *The (cross-)Wigner distribution W of $f, g \in L^2(\mathbb{R}^n)$ is the Fourier transform of the ambiguity function of f and g, i.e,*

$$W(f,g)(\xi, x) = \widehat{A(f,g)}(\xi, x) \tag{4.8}$$

$$= \int\int A(f,g)(p,q)e^{-2\pi i(p\xi+qx)}\,dpdq$$
$$= \int e^{-2\pi ip\xi} f(x+\frac{p}{2})\overline{g(x-\frac{p}{2})}\,dp\ .$$

By similar consideration as the ambiguity function, the Wigner distribution extends from a bilinear map $W : L^2(\mathbb{R}^n)\times L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^{2n})$ to a unitary map $\widetilde{W}$ of $L^2(\mathbb{R}^{2n})$ onto itself which is also a continuous bijection of $\mathcal{S}(\mathbb{R}^{2n})$ and $\mathcal{S}'(\mathbb{R}^{2n})$ onto themselves. Analogous of these facts transfer back to W on $L^2(\mathbb{R}^n)\times L^2(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)\times\mathcal{S}(\mathbb{R}^n)$ and $\mathcal{S}'(\mathbb{R}^n)\times\mathcal{S}'(\mathbb{R}^n)$.

By elementary calculations, one can obtain the following useful facts about the Wigner distribution [13].

Proposition 4.1.3 *Let $f,g\in L^2(\mathbb{R}^n)$, $a,b,c,d\in\mathbb{R}^n$ and ρ be defined as in (4.5). Then*

(i) $W(f,g)\in L^2(\mathbb{R}^{2n})$ *with* $\|W(f,g)\|_{L^2} = \|f\|_{L^2}\|g\|_{L^2}$.

(ii) $W(f,g)\in C_0(\mathbb{R}^{2n})$ *and* $\|W(f,g)\|_{L^\infty} \le \|f\|_{L^2}\|g\|_{L^2}$.

(iii) $W(f,g) = \overline{W(g,f)}$.

(iv) $W(\widehat{f},\widehat{g})(\xi,x) = W(f,g)(x,-\xi)$.

(v) $W(\rho(a,b)f,\rho(c,d)g)(\xi,x)$
$= e^{\pi i(bc-ad)}e^{2\pi i((a-c)\xi+(b-d)x)}W(f,g)(\xi-(b+d)/2, x+(a+c)/2)$.

(vi) (Moyal's Identity) $\langle W(f_1,g_1),W(f_2,g_2)\rangle = \langle f_1,f_2\rangle\langle g_2,g_1\rangle$.

4.1.4 Weyl Correspondence

Definition 4.1.4 *Let $\sigma\in\mathcal{S}'(\mathbb{R}^{2n})$ and $f,g\in\mathcal{S}(\mathbb{R}^n)$, the pseudodifferential operator $L_\sigma : \mathcal{S}(\mathbb{R}^n)\to\mathcal{S}'(\mathbb{R}^n)$ is defined implicitly by the following relation:*

$$\langle L_\sigma f,g\rangle \;=\; \langle\sigma, W(g,f)\rangle \tag{4.9}$$

$$= \langle \sigma, \overline{W(f,g)} \rangle$$
$$= \int\int \sigma(\xi,x) W(f,g)(\xi,x)\, d\xi dx$$

The distribution σ is called the symbol of the operator L_σ and L_σ is called the *Weyl transform* of σ. The *Weyl correspondence* is the one to one map between σ and the operator L_σ. Explicitly, the kernel form of the operator L_σ is given by

$$L_\sigma f(x) = \int\int \sigma(\xi, \frac{x+y}{2}) e^{2\pi i (x-y)\xi} f(y)\, dy d\xi \ . \tag{4.10}$$

We will also use the notation $\sigma(D,X)$ to denote the operator L_σ.

Of course, operator L_σ arising from distributional symbol $\sigma \in \mathcal{S}(\mathbb{R}^{2n})$ will be defined *a priori* only on the Schwartz class of function and will take values in the space of tempered distributions $\mathcal{S}'(\mathbb{R}^n)$. It is therefore natural to ask when a symbol is associated with a bounded operator on $L^2(\mathbb{R}^n)$. The following theorem summarizes some known facts along these lines and the proof can be found in [13], [31], [18].

Theorem 4.1.5 *Given $1 \le p \le 2$, let p' satisfy $\frac{1}{p} + \frac{1}{p'} = 1$. Then the Weyl correspondence is a continuous mapping of symbol $\sigma \in L^p(\mathbb{R}^{2n})$ to operators $L_\sigma \in \mathcal{I}_{p'}$, i.e, there exists a constant C_p so that*

$$||L_\sigma||_{\mathcal{I}_{p'}} \le C_p ||\sigma||_{L^p} \ , \qquad \forall \sigma \in L^p(\mathbb{R}^{2n}),$$

Moreover, for $p = 2$ the Weyl correspondence is a unitary bijection of $L^2(\mathbb{R}^{2n})$ onto $\mathcal{I}_2$. In particular,

$$||L_\sigma||_{\mathcal{I}_2} = ||\sigma||_{L^2} \ , \qquad \forall \sigma \in L^2(\mathbb{R}^{2n}) \ . \tag{4.11}$$

Remarks: In the theory of pseudodifferential operators one customarily associates to a symbol $\sigma(\xi, x)$ the operator $\sigma(D,X)_{KN}$ ("KN" for "Kohn-Nirenberg") defined by

$$\sigma(D,X)_{KN} f(x) = \int \sigma(\xi,x) e^{2\pi i x\xi} \hat{f}(\xi)\, d\xi \tag{4.12}$$
$$= \int\int \sigma(\xi,x) e^{2\pi i (x-y)\xi} f(y)\, dy d\xi \ ,$$

which bears an evident resemblance to (4.10). Just with the Weyl correspondence, one can verify the map $\sigma \to \sigma(D,X)_{KN}$ is an isomorphism from $\mathcal{S}'(\mathbb{R}^n)$ to the space of continuous linear maps from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}'(\mathbb{R}^n)$. The distributional kernel of $\sigma(D,X)_{KN}$ is $K(x,y) = (\mathcal{F}_1\sigma)(y-x,x)$, where $\mathcal{F}_j$ denotes the Fourier transform in the j-th variable.

The motivation for the definition (4.12) is that if σ is a polynomial in ξ, say $\sigma(\xi,x) = \sum_{|\alpha|\le k} a_\alpha(x)\xi^\alpha$, then

$$\sigma(D,X)_{KN}f(x) = \sum_{|\alpha|\le k} a_\alpha(x)D^\alpha f(x) \ , \tag{4.13}$$

so one obtains differential operators written in the usual way, with the differentiations on the right.

For the Weyl correspondence with the same polynomial $\sigma(\xi,x)$,

$$\sigma(D,X)f(x) = \sum_{|\alpha|\le k}\sum_{\beta\le\alpha} \frac{\alpha!}{\beta!(\alpha-\beta)!2^{|\alpha-\beta|}} D^{\alpha-\beta}a_\alpha(x)D^\beta f(x) \ , \tag{4.14}$$

(Here "$\beta \le \alpha$" means that $\beta_j \le \alpha_j$ for all j). The relation between the Weyl and Kohn-Nirenberg symbols for an operator, which derived by Moyal [29], is state in the following proposition:

Proposition 4.1.6 *If $\sigma \in \mathcal{S}'(\mathbb{R}^{2n})$, then $\sigma(D,X)_{KN} = (T\sigma)(D,X)$ where*

$$\widehat{T\sigma}(p,q) = e^{-\pi ipq}\widehat{\sigma}(p,q) \ .$$

4.1.5 Gabor Frames

A subset $\Lambda \subset \mathbb{R}^{2n}$ is a *rectangular lattice* if it has the form $\Lambda = a_1\mathbb{Z}\times\cdots\times a_{2n}\mathbb{Z}$ with $a_i > 0$. The *density* $d(\Lambda)$ of Λ is given by

$$d(\Lambda) = \frac{1}{a_1\cdots a_{2n}} \ .$$

The lattice is *square* if $a_1 = \cdots = a_{2n}$.

If $g \in L^2(\mathbb{R}^n)$, $\Lambda \subset \mathbb{R}^{2n}$ is a rectangular lattice and ρ is defined as in (4.5), then the collection of functions $\{\rho(\alpha)g\}_{\alpha\in\Lambda}$ is called the *Gabor system* generated by g and Λ, which is a collection of time-frequency shifts of g along Λ.

Gabor's fundamental work [14] proposed using a Gabor system generated by the Gaussian function

$$\phi(x) = 2^{n/4}e^{-\pi x^2}\ , \tag{4.15}$$

and a lattice Λ with density $d(\Lambda) = 1$. This system is complete in $L^2(\mathbb{R}^n)$ but it is not a frame. Seip and Wallstén [34] [35], established, for one-dimensional case $n = 1$, that the Gaussian function ϕ will generate a Gabor frame for $L^2(\mathbb{R})$ for any rectangular lattice Λ on $\mathbb{R}^2$ with density $d(\Lambda) > 1$. There is a simple extension on $\mathbb{R}^n$.

Theorem 4.1.7 *If Λ is a rectangular lattice in $\mathbb{R}^{2n}$ with density $d(\Lambda) > 1$, then the Gabor system $\{\rho(\alpha)\phi\}_{\alpha\in\Lambda}$ is a frame for $L^2(\mathbb{R}^n)$.*

Proof: Let Λ_k be rectangular lattices on $\mathbb{R}^2$ with $d(\Lambda_k) > 1$ such that $\Lambda = \Lambda_1 \times \cdots \times \Lambda_n$. Let ρ_1 be defined as in (4.5) with $n = 1$ and ϕ be one-dimensional Gaussian function. By using the one-dimensional result of Seip and Wallstén, $\{\rho_1(\lambda)\phi\}_{\lambda\in\Lambda_k}$ are frames for $L^2(\mathbb{R})$.

By using the fact that $L^2(\mathbb{R}^{m+n}) = L^2(\mathbb{R}^m) \otimes L^2(\mathbb{R}^n)$ and (viii) of Proposition 1.5.1,

$$\Big\{\rho_1(\lambda_1)\phi \otimes \cdots \otimes \rho_1(\lambda_n)\phi\Big\}_{(\lambda_1,\cdots,\lambda_n)\in\Lambda_1\times\cdots\times\Lambda_n} \tag{4.16}$$

is a frame for $L^2(\mathbb{R}^n)$, and (4.16) is the desired Gabor system. □

Lemma 4.1.8 *Let Λ be a rectangular lattice in $\mathbb{R}^{2n}$ with density $d(\Lambda) > 1$. Let $\phi_\alpha = \rho(\alpha)\phi$ and A_Λ, B_Λ to be the frame bounds for the frame $\{\phi_\alpha\}_{\alpha\in\Lambda}$. We further let $\{\widetilde{\phi}_\alpha\}_{\alpha\in\Lambda}$ to be the dual frame of $\{\phi_\alpha\}_{\alpha\in\Lambda}$. If we define*

$$\Phi_{\alpha,\beta} = W(\phi_\alpha, \phi_\beta) \quad \textit{and} \quad \widetilde{\Phi}_{\alpha,\beta} = W(\widetilde{\phi}_\alpha, \widetilde{\phi}_\beta)$$

and set $\Gamma = \Lambda \times \Lambda$. Then $\{\Phi_{\alpha,\beta}\}_{(\alpha,\beta)\in\Gamma}$ is a frame for $L^2(\mathbb{R}^{2n})$ with frame bounds A_Λ^2, B_Λ^2, dual frame $\{\widetilde{\Phi}_{\alpha,\beta}\}_{(\alpha,\beta)\in\Gamma}$, and dual frame bounds B_Λ^{-2} , A_Λ^{-2}.

Proof: First, note that the dual frame $\{\widetilde{\phi}_\alpha\}_{\alpha\in\Lambda}$ of the Gabor frame $\{\phi_\alpha\}_{\alpha\in\Lambda}$ is itself a Gabor frame using the same lattice but a different generating function $\widetilde{\phi} = S^{-1}\phi \in L^2(\mathbb{R}^n)$, where S is the frame operator for $\{\phi_\alpha\}_{\alpha\in\Lambda}$ (see Chapter 1 for the definition of frame operator). Using the same fact as in the proof of Theorem 4.1.7, $\{\phi_\alpha(x)\overline{\phi_\beta(y)}\}_{(\alpha,\beta)\in\Gamma}$ forms a frame for $L^2(\mathbb{R}^{2n}) = L^2(\mathbb{R}^n)\otimes L^2(\mathbb{R}^n)$, where $\Gamma = \Lambda\times\Lambda$. By the discussion after Definition 4.1.1, we have $W(\phi_\alpha,\phi_\beta) = \widetilde{W}(\phi_\alpha(x)\overline{\phi_\beta(y)})$, where $\widetilde{W}$ is a unitary mapping of $L^2(\mathbb{R}^{2n})$ onto itself. Since frames are preserved by unitary mappings, then $\{\Phi_{\alpha,\beta}\}_{(\alpha,\beta)\in\Gamma}$ is a frame for $L^2(\mathbb{R}^{2n})$ with dual frame $\{\widetilde{\Phi}_{\alpha,\beta}\}_{(\alpha,\beta)\in\Gamma}$. The frame bounds of $\{\Phi_{\alpha,\beta}\}_{(\alpha,\beta)\in\Gamma}$ and $\{\widetilde{\Phi}_{\alpha,\beta}\}_{(\alpha,\beta)\in\Gamma}$ follows from (v), (viii) of Proposition 1.5.1. $\square$

For later use, we need

Proposition 4.1.9

$$\Phi(\xi,x) = \Phi_{0,0}(\xi,x) = W(\phi,\phi)(\xi,x) = 2^n e^{-2\pi(\xi^2+x^2)} \tag{4.17}$$

Proof of Proposition 4.1.9: By definition, we have

$$\begin{aligned}
\Phi(\xi,x) &= W(\phi,\phi)(\xi,x)\\
&= \int e^{-2\pi ip\xi}\phi(x+\frac{p}{2})\overline{\phi(x-\frac{p}{2})}\,dp\\
&= \int e^{-2\pi ip\xi}\cdot 2^{n/4}e^{-\pi(x+\frac{p}{2})^2}\cdot 2^{n/4}e^{-\pi(x-\frac{p}{2})^2}\,dp\\
&= 2^{n/2}\int e^{-2\pi ip\xi}\cdot e^{-\pi(2x^2+\frac{p^2}{2})}\,dp\\
&= 2^n e^{-2\pi x^2}\cdot 2^{-n/2}\int e^{-2\pi ip\xi-\pi\frac{p^2}{2}}\,dp\,.
\end{aligned}$$

To prove (4.17), we need to show that

$$2^{-n/2}\int e^{-2\pi ip\xi-\pi\frac{p^2}{2}}\,dp = e^{-2\pi\xi^2}\,.$$

This can be done by completing the square, change of variable and use the fact that

$$\int_{\mathbb{R}^n} e^{-x^2/2}dx = (2\pi)^{n/2}\,. \ \square$$

Let $\alpha = (\alpha_1, \alpha_2)$, $\beta = (\beta_1, \beta_2) \in \mathbb{R}^n \times \mathbb{R}^n$, we define the linear transformation $M : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \to \mathbb{R}^{4n}$ by the formula

$$M(\alpha, \beta) = M(\alpha_1, \alpha_2, \beta_1, \beta_2) = (-\frac{\alpha_2 + \beta_2}{2}, \frac{\alpha_1 + \beta_1}{2}, \alpha_1 - \beta_1, \alpha_2 - \beta_2) . \quad (4.18)$$

Then by direct computation, we obtain

$$W(\rho(\alpha)f, \rho(\beta)g) = \rho(M(\alpha, \beta))W(f, g) , \quad (4.19)$$

where the domain of ρ on the left is $\mathbb{R}^{2n}$ whereas the domain of ρ on the right is $\mathbb{R}^{4n}$.

Note that by using equation (4.19) and the notations in Lemma 4.1.8, we obtain

$$\Phi_{\alpha,\beta} = W(\phi_\alpha, \phi_\beta) = \rho(M(\alpha, \beta))\Phi , \quad (4.20)$$

while ϕ_α denote the time-frequency shift of $\phi \in L^2(\mathbb{R}^n)$ by $\alpha \in \mathbb{R}^{2n}$ and the notation $\Phi_{\alpha,\beta}$ describe a time-frequency shift of Φ by $M(\alpha, \beta) \in \mathbb{R}^{4n}$.

4.2 Singular Values of L_σ

Our main task in this section is to estimate the decay of the singular values of the Weyl correspondence and the main theorem is state in the following. Since $\sigma \in L^2(\mathbb{R}^{2n})$, L_σ is a compact operator. Hence, we can consider the singular values $s_k(L_\sigma)$ of L_σ.

Theorem 4.2.1 *If $\sigma \in L^2(\mathbb{R}^{2n})$, then there exists constants $\epsilon, C_1, C_2 > 0$ such that the singular values of the Weyl transform L_σ satisfy*

$$\sum_{k>N} s_k(L_\sigma)^2 \leq C_1 S_\epsilon(C_2 N^{1/2n}) \ , \tag{4.21}$$

for every $N > 0$, where

$$S_\epsilon(r) = \|\sigma \cdot \chi_{B_r^C}\|_{L^2}^2 + \|\widehat{\sigma} \cdot \chi_{B_r^C}\|_{L^2}^2 + e^{-\epsilon r^2}\left(\|\sigma \cdot \chi_{B_r}\|_{L^2}^2 + \|\widehat{\sigma} \cdot \chi_{B_r}\|_{L^2}^2\right) \ . \tag{4.22}$$

In particular,

$$s_{2k}(L_\sigma)^2 \leq \frac{C_1 S_\epsilon(C_2 k^{1/2n})}{k} \ . \tag{4.23}$$

To prove Theorem 4.2.1, we have to get the approximation of the symbols σ. Given any symbols $\sigma \in L^2(\mathbb{R}^{2n})$, we can use Proposition 1.5.1 and Lemma 4.1.8 to expand σ in terms of the frame $\{\Phi_{\alpha,\beta}\}_{(\alpha,\beta)\in\Gamma}$:

$$\sigma = \sum_{(\alpha,\beta)\in\Gamma} \langle \sigma, \Phi_{\alpha,\beta}\rangle \ \widetilde{\Phi}_{\alpha,\beta} \ . \tag{4.24}$$

This series converges unconditionally in L^2-norm. Let $f, g \in L^2(\mathbb{R}^n)$, we can therefore perform the calculation on the Weyl transform L_σ:

$$\begin{aligned}
\langle L_\sigma f, g\rangle &= \langle \sigma, W(g,f)\rangle \\
&= \sum_{(\alpha,\beta)\in\Gamma} \langle \sigma, \Phi_{\alpha,\beta}\rangle \langle \widetilde{\Phi}_{\alpha,\beta}, W(g,f)\rangle \\
&= \sum_{(\alpha,\beta)\in\Gamma} \langle \sigma, \Phi_{\alpha,\beta}\rangle \langle W(\widetilde{\phi}_\alpha, \widetilde{\phi}_\beta), W(g,f)\rangle \\
&= \sum_{(\alpha,\beta)\in\Gamma} \langle \sigma, \Phi_{\alpha,\beta}\rangle \langle f, \widetilde{\phi}_\beta\rangle \langle \widetilde{\phi}_\alpha, g\rangle \ . \quad \text{(Use Moyal's identity)}
\end{aligned}$$

Hence,

$$L_\sigma f = \sum_{(\alpha,\beta)\in\Gamma} \langle \sigma, \Phi_{\alpha,\beta}\rangle \langle f, \widetilde{\phi}_\beta\rangle\ \widetilde{\phi}_\alpha\ . \tag{4.25}$$

The partial sums formed by truncating (4.24) can be used to construct finite rank approximations of L_σ. In later sections, we will design such a truncation to facilitate the estimation.

Let $B_r = \{x \in \mathbb{R}^{2n} : |x| < r\}$ and define $\Gamma_N = \Gamma \cap M^{-1}(B_N \times B_N)$, where M is defined in (4.18). Then we set

$$\sigma_N = \sum_{(\alpha,\beta)\in\Gamma_N} \langle \sigma, \Phi_{\alpha,\beta}\rangle\ \widetilde{\Phi}_{\alpha,\beta}\ . \tag{4.26}$$

Since $\{\Phi_{\alpha,\beta}\}$ is not an exact frame, (4.26) is not the frame expansion of σ_N. By a similarly calculation, the Weyl transform L_{σ_N} of σ_N is given by

$$L_{\sigma_N} f = \sum_{(\alpha,\beta)\in\Gamma_N} \langle \sigma, \Phi_{\alpha,\beta}\rangle \langle f, \widetilde{\phi}_\beta\rangle\ \widetilde{\phi}_\alpha\ . \tag{4.27}$$

Lemma 4.2.2 *There exists a constant R so that for every N, we have*

$$\operatorname{rank}(L_{\sigma_N}) \le RN^{2n}\ . \tag{4.28}$$

Proof: Let $\#$ denote the counting measure and $|\cdot|$ denote the Lebesgue measure. From (4.27), we have $\operatorname{rank}(L_{\sigma_N}) \le \#\{\alpha \in \mathbb{R}^{2n} : (\alpha,\beta) \in \Gamma_N\}$. By definition, $\Gamma_N \subset M^{-1}(B_N \times B_N) \subset B_{2N} \times B_{2N}$, so $\{\alpha \in \mathbb{R}^{2n} : (\alpha,\beta) \in \Gamma_N\} \subset \Lambda \cap B_{2N}$. The result then follows from the observation that

$$\lim_{N\to\infty} \frac{\#(\Lambda \cap B_{2N})}{|B_{2N}|} = d(\Lambda)\ ,$$

and the fact that the volume of B_{2N} is $|B_{2N}| = (2N)^{2n}|B_1|$. $\Box$

By (4.3), the singular values of L_σ are controlled by the error between L_σ and L_{σ_N} in the $\mathcal{I}_2$-norm.

Lemma 4.2.3

$$\sum_{k>\operatorname{rank}(L_{\sigma_N})} s_k(L_\sigma)^2 \le A_\Lambda^{-2} \sum_{(\alpha,\beta)\in\Gamma\setminus\Gamma_N} |\langle \sigma, \Phi_{\alpha,\beta}\rangle|^2\ . \tag{4.29}$$

Proof: Note that $\sigma - \sigma_N = \sum_{(\alpha,\beta)\in\Gamma\backslash\Gamma_N} \langle\sigma, \Phi_{\alpha,\beta}\rangle\ \widetilde{\Phi}_{\alpha,\beta}$ and $\{\widetilde{\Phi}_{\alpha,\beta}\}$ has frame bounds $B_\Lambda^{-2}, A_\Lambda^{-2}$. By (4.3)

$$\begin{aligned}
\sum_{k>\mathrm{rank}(L_{\sigma_N})} s_k(L_\sigma)^2 &\leq \inf\{\|L_\sigma - T\|_{\mathcal{I}_2}^2 : rank(T) \leq rank(L_{\sigma_N})\} \\
&\leq \|L_\sigma - L_{\sigma_N}\|_{\mathcal{I}_2}^2 \\
&= \|\sigma - \sigma_N\|_{L^2}^2 \quad \text{(Use Theorem 4.1.5)} \\
&= \|\sum_{(\alpha,\beta)\in\Gamma\backslash\Gamma_N} \langle\sigma, \Phi_{\alpha,\beta}\rangle\ \widetilde{\Phi}_{\alpha,\beta}\|_{L^2}^2 \\
&\leq A_\Lambda^{-2} \sum_{(\alpha,\beta)\in\Gamma\backslash\Gamma_N} |\langle\sigma, \Phi_{\alpha,\beta}\rangle|^2\ . \quad \square
\end{aligned}$$

Our main task in the following is to estimate the right hand side of (4.29). To do this, we first set up some notations. Note that if Λ is a square lattice, then

$$M(\Gamma_N) = M(\Gamma) \cap (B_N \times B_N) \ \text{ and } \ M(\Gamma) \subset \frac{1}{2}\Gamma = \frac{1}{2}\Lambda \times \frac{1}{2}\Lambda\ . \tag{4.30}$$

So,

$$\begin{aligned}
M(\Gamma\backslash\Gamma_N) &\subset \frac{1}{2}\Lambda \cap (B_N \times B_N)^C \\
&= (\frac{1}{2}\Lambda \cap B_N^C) \times (\frac{1}{2}\Lambda \cap B_N^C)\ .
\end{aligned}$$

We also have

$$\Phi(\xi, x) = 2^n e^{-2\pi(\xi^2+x^2)} \ \text{ and } \ \widehat{\Phi}(p,q) = e^{-\pi(p^2+q^2)/2}\ .$$

Define

$$G(\xi, x) = 2^{n/2} e^{-\pi(\xi^2+x^2)} \ \text{ and } \ H(p,q) = e^{-\pi(p^2+q^2)/2}\ ,$$

so that $G^2 = \Phi$ and $H^2 = \hat{\Phi}$. If we let τ_β to be the translation operator such that $\tau_\beta f(\alpha) = f(\alpha - \beta)$. By direct computation, we obtain

$$\rho(\mu,\nu)\Phi = \rho(\mu,\nu)G \cdot \tau_{-\mu}G \ \text{ and } \ \rho(\nu,\mu)\widehat{\Phi} = \rho(\nu,\mu)H \cdot \tau_{-\nu}H.$$

Moreover, a Gabor system generated by any Gaussian function on any arbitrary rectangular lattice always has a upper frame bound. Hence, the sequences

$\{\rho(\mu,\nu)G\}_{(\mu,\nu)\in\frac{1}{2}\Gamma}$ and $\{\rho(\nu,\mu)H\}_{(\mu,\nu)\in\frac{1}{2}\Gamma}$ would have upper frame bounds B_G and B_H respectively.

Lemma 4.2.4

$$\sum_{(\alpha,\beta)\in\Gamma\setminus\Gamma_N} |\langle\sigma,\Phi_{\alpha,\beta}\rangle|^2 \le C\Big(\sum_{\mu\in\frac{1}{2}\Lambda\cap B_N^C} \|\sigma\cdot\tau_{-\mu}\|_{L^2}^2 + \sum_{\nu\in\frac{1}{2}\Lambda\cap B_N^C} \|\widehat{\sigma}\cdot\tau_\nu H\|_{L^2}^2\Big). \tag{4.31}$$

Proof:

$$\begin{aligned}
&\sum_{(\alpha,\beta)\in\Gamma\setminus\Gamma_N} |\langle\sigma,\Phi_{\alpha,\beta}\rangle|^2 && (4.32)\\
&= \sum_{(\alpha,\beta)\in\Gamma\setminus\Gamma_N} |\langle\sigma,\rho(M(\alpha,\beta))\Phi\rangle|^2\\
&= \sum_{(\mu,\nu)\in M(\Gamma\setminus\Gamma_N)} |\langle\sigma,\rho(\mu,\nu)\Phi\rangle|^2\\
&\le \sum_{\mu\in\frac{1}{2}\Lambda\cap B_N^C}\sum_{\nu\in\frac{1}{2}\Lambda} |\langle\sigma,\rho(\mu,\nu)\Phi\rangle|^2 + \sum_{\nu\in\frac{1}{2}\Lambda\cap B_N^C}\sum_{\mu\in\frac{1}{2}\Lambda} |\langle\sigma,\rho(\mu,\nu)\Phi\rangle|^2\,. && (4.33)
\end{aligned}$$

We now estimate each of the sums in (4.33). For the first sum, we have

$$\begin{aligned}
\sum_{\mu\in\frac{1}{2}\Lambda\cap B_N^C}\sum_{\nu\in\frac{1}{2}\Lambda} |\langle\sigma,\rho(\mu,\nu)\Phi\rangle|^2 &= \sum_{\mu\in\frac{1}{2}\Lambda\cap B_N^C}\sum_{\nu\in\frac{1}{2}\Lambda} |\langle\sigma\cdot\tau_{-\mu}G,\rho(\mu,\nu)G\rangle|^2\\
&\le B_G \sum_{\mu\in\frac{1}{2}\Lambda\cap B_N^C} \|\sigma\cdot\tau_{-\mu}G\|_{L^2}^2\,.
\end{aligned}$$

For the second sum,

$$\begin{aligned}
\sum_{\nu\in\frac{1}{2}\Lambda\cap B_N^C}\sum_{\mu\in\frac{1}{2}\Lambda} |\langle\sigma,\rho(\mu,\nu)\Phi\rangle|^2 &= \sum_{\nu\in\frac{1}{2}\Lambda\cap B_N^C}\sum_{\mu\in\frac{1}{2}\Lambda} |\langle\widehat{\sigma},\rho(-\nu,\mu)\widehat{\Phi}\langle|^2\\
&= \sum_{\nu\in\frac{1}{2}\Lambda\cap B_N^C}\sum_{\mu\in\frac{1}{2}\Lambda} |\langle\widehat{\sigma}\cdot\tau_\nu H,\rho(-\nu,\mu)H\rangle|^2\\
&\le B_H \sum_{\nu\in\frac{1}{2}\Lambda\cap B_N^C} \|\widehat{\sigma}\cdot\tau_\nu H\|_{L^2}^2\,. \quad \square
\end{aligned}$$

Lemma 4.2.5 *There exists constants* $\epsilon, C_1, C_2 > 0$ *such that*

(i) $\sum_{\mu\in\frac{1}{2}\Lambda\cap B_N^C} \|\sigma\cdot\tau_{-\mu}G\|_{L^2}^2 \le C_1\|\sigma\cdot\chi_{B_{N/2}^C}\|_{L^2}^2 + C_2 e^{-\epsilon N^2}\|\sigma\cdot\chi_{B_{N/2}}\|_{L^2}^2$,

(ii) $\sum_{\nu\in\frac{1}{2}\Lambda\cap B_N^C}\|\widehat{\sigma}\cdot\tau_\nu H\|_{L^2}^2 \le C_1\|\widehat{\sigma}\cdot\chi_{B^C_{B_{N/2}}}\|_{L^2}^2 + C_2 e^{-\epsilon N^2}\|\widehat{\sigma}\cdot\chi_{B_{N/2}}\|_{L^2}^2$.

Proof: We prove only (i) and (ii) is similar. Define

$$G_N = \sum_{\mu\in\frac{1}{2}\Lambda\cap B_N^C}(\tau_{-\mu}G)^2 = \sum_{\mu\in\frac{1}{2}\Lambda\cap B_N^C}\tau_{-\mu}\Phi\ .$$

By Tonelli's Theorem [32],

$$\sum_{\mu\in\frac{1}{2}\Lambda\cap B_N^C}\|\sigma\cdot\tau_{-\mu}G\|_{L^2}^2 = \int\int|\sigma(\xi,x)|^2 G_N(\xi,x)\ d\xi dx\ .$$

Since G is a Gaussian function, we have $C_1 = \sup\|G_N\|_{L^\infty} < \infty$. Therefore,

$$\int\int_{B^C_{N/2}}|\sigma(\xi,x)|^2 G_N(\xi,x)\ d\xi dx \le C_1\|\sigma\cdot\chi_{B^C_{N/2}}\|_{L^2}^2\ .$$

Further,

$$\int\int_{B_{N/2}}|\sigma(\xi,x)|^2 G_N(\xi,x)\ d\xi dx \le \|G_N\cdot\chi_{B_{N/2}}\|_{L^\infty}\|\sigma\cdot\chi_{B_{N/2}}\|_{L^2}^2.$$

However, if $(\xi,x)\in B_{N/2}$ and $|\mu|\ge N$, then $|(\xi,x)-\mu|\ge|\mu|-N/2$, so

$$\begin{aligned} G_N(\xi,x) &= \sum_{\mu\in\frac{1}{2}\Lambda\cap B_N^C} 2^n e^{-2\pi((\xi,x)-\mu)^2} \\ &\le 2^n\sum_{\mu\in\frac{1}{2}\Lambda\cap B_N^C} e^{-2\pi(|\mu|-N/2)^2} \\ &\le C_2\int_{|x|\ge N/2} e^{-2\pi x^2}dx \\ &\le C_2 e^{-(\pi n/4)N^2}. \end{aligned}$$

By combining all this, the result follows immediately. □

We now combine all this lemmas to complete the proof of Theorem 4.2.1.

Proof of Theorem 4.2.1: By combining (4.29) (4.33) and Lamma 4.2.5, we see that there exist constants ϵ and C so that

$$\sum_{k>\mathrm{rank}(L_{\sigma_N})} s_k(L_\sigma)^2 \le CS_\epsilon(N/2)\ ,$$

where $S_\epsilon(r)$ is defined in (4.22). By Lemma 4.2.2, there exists constant R so that $\mathrm{rank}(L_{\sigma_N}) \leq RN^{2n}$ for every N. Therefore

$$\sum_{k>RN^{2n}} s_k(L_\sigma)^2 \leq CS_\epsilon(N/2) .$$

By reparametrizing this equation, we obtain (4.21). Finally, (4.23) follows from (4.21) because singular values are arranged in decreasing order, so

$$Ns_{2N}(L_\sigma)^2 \leq \sum_{k=N+1}^{2N} s_k(L_\sigma)^2 \leq C_1 S_\epsilon(C_2 N^{1/2n}) . \quad \square$$

4.3 The Calderón-Vaillancourt Theorem

4.3.1 Hölder-Zygmund Spaces

Let $k \in \mathbb{N}$, then

$$C^k(\mathbb{R}^n) = \left\{ f \in C(\mathbb{R}^n) : D^\alpha f \in C(\mathbb{R}) \text{ if } |\alpha| \leq k \right\} \tag{4.34}$$

are Banach spaces equipped with the norm

$$\|f\|_{C^k} = \sum_{|\alpha| \leq k} \|D^\alpha f\|_\infty . \tag{4.35}$$

The notation D^α is defined as in (4.4) and $|\alpha| = \sum_{j=1}^n \alpha_j$.

The first step beyond the space $C^k(\mathbb{R}^n)$ with $k \in \mathbb{N}$ are the Hölder space $C^s(\mathbb{R}^n)$ with $0 < s \neq$ integer. The spaces fill the gaps between the spaces $C^k(\mathbb{R}^n)$ with $k \in \mathbb{N}$.

Let $0 < \sigma < 1$, then we introduce the norm

$$\|f\|_{C^\sigma} = \|f\|_\infty + \sup \frac{|f(x) - f(y)|}{|x-y|^\sigma} , \tag{4.36}$$

where the supremum is taken over all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ with $x \neq y$. Let $s \in \mathbb{R}$, then we put

$$s = [s] + \{s\} , \tag{4.37}$$

where $[s]$ is integer, whereas $0 \leq \{s\} < 1$.

Definition 4.3.1 *Let $0 < s \neq$ integer, then the Hölder-Zygmund spaces $\Lambda^s(\mathbb{R}^n)$ are defined by*

$$\Lambda^s(\mathbb{R}^n) = \left\{ f \in C(\mathbb{R}^n) : \|f\|_{\Lambda^s} = \|f\|_{C^{[s]}} + \sum_{|\alpha|=[s]} \|D^\alpha f\|_{C^{\{s\}}} < \infty \right\}. \tag{4.38}$$

4.3.2 Smooth Dyadic Resolution of Unity

Definition 4.3.2 *We call $\{\varphi_j\}_{j=0}^{\infty} \subset C^\infty(\mathbb{R}^n)$ a smooth dyadic resolution of unity in $\mathbb{R}^n$ if φ_j satisfy*

$$\begin{cases} \mathrm{supp}\varphi_0 \subset \{\xi \in \mathbb{R}^n : |\xi| \leq 2\}, \\ \mathrm{supp}\varphi_k \subset \{\xi \in \mathbb{R}^n : 2^{k-1} \leq |\xi| \leq 2^{k+1}\}, \quad k \in \mathbb{N}, \end{cases} \tag{4.39}$$

$$\sum_{j=0}^{\infty} \varphi_j(\xi) = 1, \quad \text{if } \xi \in \mathbb{R}^n \tag{4.40}$$

and for any multi-index α,

$$\sup_{\xi \in \mathbb{R}^n,\ j \in \mathbb{N}_0} 2^{j|\alpha|} |D^\alpha \varphi_j(\xi)| < \infty \tag{4.41}$$

Example: Let φ_0 be an C^∞ function on $\mathbb{R}^n$ with $\mathrm{supp}\varphi_0 \subset \{\xi \in \mathbb{R}^n : |\xi| \leq 2\}$ and $\varphi_0(\xi) = 1$ if $|\xi| \leq 1$. Then $\{\varphi_j\}_{j=0}^{\infty}$ defined by

$$\varphi_j(\xi) = \varphi_0(2^{-j}\xi) - \varphi_0(2^{-j+1}\xi), \qquad j \in \mathbb{N}, \tag{4.42}$$

is a smooth dyadic resolution of unity.

Let $\varphi \in C(\mathbb{R}^n)$, the operator $\mathcal{K}_\varphi$ is defined by

$$\mathcal{K}_\varphi f = (\varphi \widehat{f})^\vee. \tag{4.43}$$

If $\varphi \in C_c^\infty$, (4.43) make sense for any $f \in \mathcal{S}'$.

By [40], we have the following characterization of the Hölder-Zygmund classes. Let $\Phi = \{\varphi_j\}_{j=0}^{\infty}$ be a smooth dyadic resolution of unity in $\mathbb{R}^n$ and let $\mathcal{K}_\varphi$ be defined as in (4.43),

$$\Lambda^s(\mathbb{R}^n) = \left\{ f \in \mathcal{S}' : \sup_{j \geq 0} 2^{sj} \|\mathcal{K}_{\varphi_j} f\|_\infty < \infty \right\} \tag{4.44}$$

in the sense of equivalent norms.

4.3.3 The proof of the Calderón-Vaillancourt Theorem

The Calderón-Vaillancourt Theorem is states as follows:

Theorem 4.3.3 *If $\sigma \in \Lambda^s(\mathbb{R}^{2n})$ with $s > 2n$, then L_σ is a bounded operator on $L^2(\mathbb{R}^n)$.*

The proof of this theorem use Gabor frame expansions in a different manner than previous sections. We still want to approximate the symbol σ by symbol σ_N. In order that σ_N share the smoothness properties of σ yet be an element of $L^2(\mathbb{R}^n)$, we choose a smooth, compactly supported, non-negative function $m \in C_c^\infty(\mathbb{R}^{2n})$ satisfying $m(\xi, x) = 1$ if $|(\xi, x)| \leq 1$ and define,

$$\sigma_N(\xi, x) = m(\frac{\xi}{N}, \frac{x}{N})\sigma(\xi, x). \tag{4.45}$$

Note that by definition, σ_N converges to σ uniformly on compact sets.

Lemma 4.3.4 *Let $\tau_\beta f(\cdot) = f(\cdot - \beta)$ be the translation operator as before and let $\eta = \left(\frac{\alpha_2+\beta_2}{2}, -\frac{(\alpha_1+\beta_1)}{2}\right)$. Then*

$$|\langle \sigma_N, \Phi_{\alpha,\beta} \rangle| = |(\widehat{\sigma_N \cdot \tau_\eta} \Phi)(\alpha - \beta)| \tag{4.46}$$

Proof: By (4.18) and (4.20), we can compute as follows:

$$\begin{aligned}
\langle \sigma_N, \Phi_{\alpha,\beta} \rangle &= \langle \sigma_N, \rho(M(\alpha, \beta))\Phi \rangle \\
&= C \int\int \sigma_N(\xi, x) e^{-2\pi i((\alpha_1-\beta_1)\xi + (\alpha_2-\beta_2)x)} \\
&\quad \times \Phi\Big(\xi - \frac{\alpha_2+\beta_2}{2},\ x + \frac{\alpha_1+\beta_1}{2}\Big)\ d\xi dx \\
&= C \int\int \sigma_N(\xi, x) e^{-2\pi i(\alpha_1-\beta_1)(\xi,x)} \tau_\eta \Phi(\xi, x)\ d\xi dx \\
&= C(\widehat{\sigma_N \cdot \tau_\eta} \Phi)(\alpha - \beta)\ . \ \square
\end{aligned}$$

The key to the proof of the Calderón-Vaillancourt Theorem is to show that there is a sequence $k = \{k(\alpha)\} \in l^1(\Lambda)$, independent of N and such that $|\langle \sigma_N, \Phi_{\alpha,\beta} \rangle| \leq k(\alpha - \beta)$, then we could use the uniform pointwise convergence of

σ_N to σ on compact sets to derive a weak convergence of L_{σ_N} to L_σ. Hence, L_σ is bounded on $L^2(\mathbb{R}^n)$.

Proposition 4.3.5 *Let $\sigma \in \mathcal{S}'(\mathbb{R}^{2n})$ and $k = \{k(\alpha)\} \in l^1(\Lambda)$ be a nonnegative sequence. Assume that $\sigma_N \in L^2(\mathbb{R}^{2n})$ are such that*

$$|\langle \sigma_N, \Phi_{\alpha,\beta}\rangle| \leq k(\alpha - \beta) \quad \forall \alpha, \beta \in \Lambda \tag{4.47}$$

and suppose that $L_{\sigma_N} \to L_\sigma$ weakly as operators from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}'(\mathbb{R}^n)$, i.e. $\langle L_{\sigma_N} f, g\rangle \to \langle L_\sigma f, g\rangle$ for all $f, g \in \mathcal{S}(\mathbb{R}^n)$. Then the operator L_σ is bounded on $L^2(\mathbb{R}^n)$ and its operator norm satisfies

$$\|L_\sigma\|_{\mathcal{B}(L^2)} \leq A_\Lambda^{-1} \|k\|_{l^1}\ , \tag{4.48}$$

where A_Λ^{-1} is the dual frame bound in Lemma 4.1.8.

Proof: Since each L_{σ_N} is a bounded operator on $L^2(\mathbb{R}^n)$. By (4.25), we have

$$\langle L_{\sigma_N} f, g\rangle = \sum_{\alpha,\beta\in\Lambda} \langle \sigma_N, \Phi_{\alpha,\beta}\rangle \langle f, \widetilde{\phi}_\beta\rangle \langle \widetilde{\phi}_\alpha, g\rangle.$$

Let $u = \{|\langle f, \widetilde{\phi}_\alpha\rangle|\}$ and $v = \{|\langle \widetilde{\phi}_\alpha, g\rangle|\}$. Then for each $f, g \in L^2(\mathbb{R}^n)$ we have

$$\begin{aligned}
|\langle L_{\sigma_N} f, g\rangle| &\leq \sum_{\alpha,\beta\in\Lambda} |\langle \sigma_N, \Phi_{\alpha,\beta}\rangle \langle f, \widetilde{\phi}_\beta\rangle \langle \widetilde{\phi}_\alpha, g\rangle| \\
&\leq \sum_{\alpha,\beta\in\Lambda} k(\alpha - \beta) u(\beta) v(\alpha) \\
&= \langle k * u, v\rangle_{l^2} \\
&\leq \|k\|_{l^1} \cdot \|u\|_{l^2} \cdot \|v\|_{l^2} \quad \text{(Young's inequality)} \\
&\leq A_\Lambda^{-1} \|k\|_{l^1} \cdot \|f\|_{L^2} \cdot \|g\|_{L^2},
\end{aligned}$$

where $\{\widetilde{\phi}_\alpha\}$ is the dual frame of $\{\phi_\alpha\}$ with frame bounds B_Λ^{-1} and A_Λ^{-1}. It follows immediately from this that L_σ is bounded on $L^2(\mathbb{R}^n)$ if $L_{\sigma_N} \to L_\sigma$ weakly as operators form $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}'(\mathbb{R}^n)$. □

Recall that the Weyl correspondence employs the Wigner distribution to define 1-1 correspondence between tempered distribution $\sigma \in \mathcal{S}'(\mathbb{R}^{2n})$ and the pseudodifferential operators $L_\sigma : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$. In particular, if $\sigma \in C(\mathbb{R}^{2n})$ and σ_N is defined by (4.45), then the σ_N are uniformly bounded in L^∞ norm and converges to σ uniformly on compact sets. Since for each $f, g \in \mathcal{S}(\mathbb{R}^n)$, we have $\langle L_\sigma f, g\rangle - \langle L_{\sigma_N} f, g\rangle = \langle \sigma - \sigma_N, W(g, f)\rangle$ with $W(g, f) \in \mathcal{S}(\mathbb{R}^{2n})$ it follows that $L_{\sigma_N} \to L_\sigma$ weakly.

Next, we want to estimate the decay of $(\widehat{\sigma_N \cdot \tau_\eta \Phi})(\alpha)$ independently of N and η. From now on, let $v_0 \in C_c^\infty$ be a function such that $\text{supp}(v_0) \subset \{\gamma \in \mathbb{R}^n : |\gamma| < 2\}$ and such that $v_0(\gamma) = 1$ if $|\gamma| \leq 1$. For each $j > 0$ define $v_j(\gamma) = v_0(2^{-j}\gamma) - v_0(2^{-(j-1)}\gamma)$ as in the example. Then $\{v_j\}_{j=0}^\infty$ is a smooth dyadic resolution of unity. For our purpose, we will impose additional restrictions on v_0, namely that $1/2 \leq v_0(\gamma) \leq 1$ when $1 \leq |\gamma| \leq 3/2$ and that $0 \leq v_0(\gamma) \leq 1/2$ when $3/2 \leq |\gamma| \leq 2$. Then since $v_j(\gamma) = v_1(2^{-(j-1)}\gamma)$ for each $j > 0$, we have

$$3 \cdot 2^{j-2} \leq |\gamma| \leq 3 \cdot 2^{j-1}\ , \quad \forall j > 0, \quad \Rightarrow \quad \frac{1}{2} \leq v_j(\gamma) \leq 1\ .$$

Now $|\gamma|$ is comparable to 2^j if $\gamma \in \text{supp}v_j$ and $v_j(\gamma)$ is comparable to 1 if γ is in the annulus $3 \cdot 2^{j-2} \leq |\gamma| \leq 3 \cdot 2^{j-1}$, which is contained in $\text{supp}v_j$.

Lemma 4.3.6 *Let $\{v_j\}_{j=0}^\infty$ be the smooth dyadic resolution of unity defined as above. Let $f \in \Lambda^s$ with $s > 0$, and let $t = s - \epsilon > 0$ with $\epsilon < 1$. Suppose $\psi \in \mathcal{S}(\mathbb{R}^n)$, then there exists a constant C, independent of j and b, such that*

$$\|\check{v}_j * (f \cdot \tau_b \psi)\|_{L^1} \leq C 2^{-tj} \tag{4.49}$$

Proof: First, note that there exists a constant C_1 independent of b such that

$$\|f \cdot \tau_b \psi\|_{\Lambda^s} \leq C_1 \|f\|_{\Lambda^s}\ .$$

Therefore, by the norm equivalence (4.44),

$$C_2 = \sup_b \sup_{j \geq 0} 2^{sj} \|\check{v}_j * (f \cdot \tau_b \psi)\|_{L^\infty} < \infty\ . \tag{4.50}$$

Our goal is to obtain a similar result with L^∞ replaced by L^1 and s replaced by $t = s - \epsilon$. Fix b and j, and define $B = B_{2^{\epsilon j/n}}(b)$, the ball of radius $2^{\epsilon j/n}$ centered at b. Then by (4.50),

$$\begin{aligned}\int_B |(\check{v}_j * (f \cdot \tau_b \psi))(x)|\, dx &\leq |B| \|\check{v}_j * (f \tau_b \cdot \psi)\|_{L^\infty} \\ &\leq |B_1| (2^{\delta j})^n \cdot C_2 2^{-sj} = C_3 2^{-tj}, \qquad (4.51)\end{aligned}$$

where B_1 is ball of radius 1.

In order to derive a similar estimate for the integral outside B, define

$$F_1 = f \cdot \tau_b \psi \cdot \chi_{\frac{1}{2}B} \quad \text{and} \quad F_2 = f \cdot \tau_b \psi \cdot (1 - \chi_{\frac{1}{2}B}),$$

so that $\check{v}_j * (f \cdot \tau_b \psi) = \check{v}_j * F_1 + \check{v}_j * F_2$. Then it suffices to show that $\|\check{v}_j * F_i\|_{L^1} \leq C 2^{-tj}$ for some constant C. Note that $\check{v}_j(x) = 2^{j-1} \check{v}_1(2^{j-1}x)$ and that ψ and $\check{v}_1$ are both Schwartz-class functions, hence decay faster than any polynomial. In particular, for each M there exists a constant $K = K(M)$ so that $|\psi(x)|$, $|\check{v}_1(x)| \leq K|x|^{-M}$. Hence,

$$|\check{v}_j(x)| \leq K 2^{j-1} |2^{j-1} x|^{-M} \leq K 2^{-(M+1)-(M+1)j} |x|^{-M}\,, \qquad \forall x \in \mathbb{R}^n.$$

Now, if $x \notin B$ and $y \in \frac{1}{2}B$, then $|x - y| \geq |x - b|/2$. Hence, for each x,

$$\begin{aligned}|(\check{v}_j * F_1)(x)| &\leq \int_{\frac{1}{2}B} |\check{v}_j(x-y) F_1(y)|\, dy \\ &\leq \|f\|_{L^\infty} \|\psi\|_{L^\infty} K 2^{-(M+1)-(M+1)j} \int_{\frac{1}{2}B} |x-y|^{-M}\, dy \\ &\leq \|f\|_{L^\infty} \|\psi\|_{L^\infty} K 2^{-(M+1)-(M+1)j} |\frac{1}{2}B| \left(\frac{|x-b|}{2}\right)^{-M} \\ &= C_4 2^{-(1+M-\epsilon)j} |x-b|^{-M},\end{aligned}$$

with C_4 depending on M, but not on b or x. Therefore, taking $M > n$ and $M > t + \epsilon - 1$, we have

$$\int_{B^C} |\check{v}_j * F_1(x)|\, dx \leq C_5 2^{-(1+M-\epsilon)j} \leq C_5 2^{-tj}, \qquad (4.52)$$

with C_5 depending only on M. Finally, it we also take $M > tn/\epsilon$, then

$$\begin{aligned}
\int_{B^C} |\check{v_j} * F_2(x)|\, dx &\leq \|\check{v_j}\|_{L^1} \cdot \|F_2 \cdot \chi_{B^C}\|_{L^\infty} \\
&\leq \|\check{v_1}\|_{L^1} \cdot \|f\|_{L^\infty} \cdot \|\tau_b\psi \cdot \chi_{B^C}\|_{L^\infty} \\
&\leq \|\check{v_1}\|_{L^1} \cdot \|f\|_{L^\infty} \cdot K \sup_{x \notin B} |x-b|^{-M} \\
&\leq C_6 2^{-\epsilon j M/n} \\
&\leq C_6 2^{-tj}\ , \qquad (4.53)
\end{aligned}$$

with C_6 depending only on M. combining (4.51), (4.52) and (4.53), we obtain (4.49). □

Lemma 4.3.7 *Suppose we have the assumptions of Lemma 4.3.6, then there exist a constant C, depending only on n, ψ and f, such that*

$$|(\widehat{f \cdot \tau_b\psi})(\gamma)| \leq C\left(\frac{4|\gamma|}{3}\right)^{-t}, \quad \forall \gamma, b \in \mathbb{R}^n. \qquad (4.54)$$

Proof: By Proposition 4.3.6, we have equation (4.49). If $\gamma \in \mathbb{R}^n$ is given, then $3 \cdot 2^{j-2} \leq |\gamma| \leq 3 \cdot 2^{j-1}$ for some j, so

$$\begin{aligned}
|(\widehat{f \cdot \tau_b\psi})(\gamma)| &\leq \frac{\|v_j \cdot (\widehat{f \cdot \tau_b\psi})\|_{L^\infty}}{v_j(\gamma)} \\
&\leq 2\|\check{v_j} * (f \cdot \tau_b\psi)\|_{L^1} \\
&\leq 2C2^{-tj} \qquad (\text{Use}(4.49)) \\
&\leq 2C\left(\frac{4|\gamma|}{3}\right)^{-t}. \ \square
\end{aligned}$$

By combining Lemma 4.3.6 and Lemma 4.3.7, we can prove Theorem 4.3.3.

Proof of Theorem 4.3.3: Assume that $\sigma \in \Lambda^s(\mathbb{R}^n)$, and let σ_N be defined by (4.45). Fix $t = s - \epsilon > 0$. Then by (4.46) and Lemma 4.3.7, there exists a constant C, independent of N, such that

$$|\langle \sigma_N, \Phi_{\alpha,\beta}\rangle| \leq C(1 + (\alpha - \beta)^2)^{-t/2}.$$

Define $k(\alpha) = (1+\alpha^2)^{-t/2}$. Then $k \in l^1(\Lambda)$ since $t > 2n$ and Λ is a rectangular lattice on $\mathbb{R}^{2n}$. Since $\sigma_N \to \sigma$ uniformly on compact sets, the Weyl transform L_{σ_N} converges weakly to L_σ. Hence the conditions of Proposition 4.3.5 are fulfilled, and therefore L_σ extends to a bounded operator on $L^2(\mathbb{R}^n)$. $\square$.

Bibliography

[1] G. Beylkin, R. coifman and V. Rokhlin, *Fast Wavelet Transforms and Numerical Algorithm I*, Comm. Pure and Appl. Math. vol.XLIV(1991), pp.141-183.

[2] G. Beylkin, *Wavelets , Multiresolution Analysis, and Fast Numerical Algorithm*, 1991, A draft of INRIA Lecture notes.

[3] G. Beylkin, *On the representation of operators in bases of compactly supported wavelets*, Vol. 6, No. 6, pp.1716-1740, December 1992.

[4] Campanato, S., *Propriétà di inclusione per spaci di Morrey*, Ricerche Mat, 12 (1963), pp.67-86.

[5] A. Cohen, *Ondelettes , analyses multirésolutions et filtres miroir en quadrature*, Ann. Inst. H. Poincaré, Anal. non linéaire, 7 , pp.439-459.

[6] John B, Conway, *A Course in Functional Analysis*, Springer-Verlag, 1990, 2nd edition.

[7] I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Comm. Pure and Appl. Math. 41(1988), pp.909-996.

[8] I. Daubechies, *Ten Lectures on Wavelets*, 1992.

[9] Guy David and Jean-Lin Journé, *A boundedness criterion for generalized Calderón-Zygmund operator*, Annals of Mathematics, 120(1984), pp.371-397.

[10] Deng Donggao, YAN Lixin, and Yang Qisiang, *Blocking Analysis and T(1) Theorem*, preprint.

[11] Fefferman, C., *Characterizations of bounded mean oscillation*, Bull. Amer. Math. Soc. 77(1971), pp.587-588.

[12] Fefferman, C. and E. M. Stein, *H^p spaces of several variables*, Acta Math. 129(1972), pp.137-193.

[13] Gerald B. Folland, *Harmonic Analysis in Phase Space*, Annals of mathematics studies, Study 122, Princeton University Press.

[14] D. Gabor, *Theory of communications*, J. Inst. Elec. Eng. (London) 93(1946), pp.429-457.

[15] I. C. Gohberg, M. G. Krein, *Introduction to the Thoery of Linear Non-selfadjoint operators*, translations of mathematical monographs, vol 18, 1969, American Mathematical Society.

[16] Han, Y. S., Hofmann, S., *T(1) theorem for Besov and Tribel-Lizorkin spaces*, Tran. Amer. Math. Soc. , 237(1993), pp.839-853.

[17] C. Heil, J. Ramanathan, P. Topiwala, *Singular Values of Compact Pseudodifferential Operators*, Journal of Functional Analysis, Vol. 150, No.2, November 1997.

[18] R. Howe, *Quantum mechanics and partial differential equations*, J. Funct. Anal. 38(1980), pp.188-254.

[19] F. John and L. Nirenberg, *On functions of bounded mean oscillation*, Comm. Pure Appl. Math 14(1961), pp.415-426.

[20] A. N. Kolmogorov, S. V. Fomin, *Introductory Real Analysis*, Revised English Edition Translated and Edited by Richard A. Silverman, pp.299

[21] W Lawton, *Tight frames of compactly supported wavelets*, J. Math. Phys., Vol31. (1990), pp.1898-1901.

[22] S. Mallat, *Multiresolution approximation and wavelets*, Trans. Am. Math. Soc. 315(1989), pp.69-88.

[23] Y. Meyer, *Principe d'incertitude, bases hilbertiennes et algébres d'opérateurs*, Séminaire Bourbaki(Astérisque), Société Mathématique de France, Astérisque, 1985-1986.

[24] Y. Meyer, *Ondelettes et fonctions slpines*, Technical report, séminaire edp, Ecole Polytechnique, Paris, France, 1986.

[25] Y. Meyer, *Wavelets and operators*, in *Analysis at Urbana*, Vol. 1, E.Berkson, N. T. Peck, and J. Uhl, eds., London maths. Soc., Lecture Notes Series 137, 1989, pp.475-487.

[26] Y. Meyer, *Wavelets and operators*, Cambridge studies in advanced mathematics 37, 1992.

[27] Y. Meyer & Ronald Coifman, *Wavelets , Calderón-Zygmund and multilinear operators*, Cambridge studies in advanced mathematics 48, 1997.

[28] Y. Meyer, *La Minimalité de le espace de Besov* B_1^{01} *et la continuité des operateurs definis par des integrals singulieres*, Monografias de matematicas, Vol.4, Univ. Autonoma de Madrid, 1986.

[29] J. E. Moyal, *Quantum mechanics as a statistical theory*, Proc. Camb. Phil. Soc. 45(1949), pp.99-124.

[30] J. Peetre, *On convolution operator leaving* $L^{p,\lambda}$ *invariant*, Ann. math. Pura Appl. 72(1966), pp.295-304.

[31] J. C. T. Pool, *Mathematical aspects of the Weyl correspondence*, J. math. Phys. 7(1966), pp.66-76.

[32] H. L. Royden, *Real Analysis*, third edition, 1988, Macmillan Publishing Company.

[33] S. Spanne, *Sur l'interpolation entre les espaces* $\mathcal{L}_k^{p,\phi}$, Ann. Scuola Norm. Sup. Pisa 20(1966), pp.625-648.

[34] K. Seip, *Density theorems for sampling and interpolation in the Bargmann-Fock space I*, J. Reine Angew. Math. 429(1992), pp.91-106.

[35] K. Seip and R. Wallstén, *Sampling and interpolation in the Bargmann-Fock space II*, J. Reine Angew. Math. 429(1992), pp.107-113.

[36] E. M. Stein, *Singular integral, harmonic functions, and differentiability properties of functions of several variables*, Proc. Symp. in Pure Math. 10(1967), pp.316-335.

[37] E. M. Stein, *Harmonic Analysis* Princeton University Press, 1993.

[38] J. O. Stromberg, *A Modified Franklin System and Higher-Order Spline Systems on* $\mathbb{R}^n$ *as Unconditional Bases for Hardy Spaces, Conference on harmonic analysis in honour of Antoni Zygmund*, Wadworth math series, 1983, pp.475-493.

[39] Hans Triebel, *Theory of Function Spaces II*, Monographs in Mathematics, 1992, Birkhäuser.

[40] Hans Triebel, *On spaces of* $B^s_{\infty,q}$ *type and* $\mathcal{C}^s$ *type*, Math. Nachr. 85(1978), pp.75-90.

[41] Yang Qi Xiang, *Fast Algorithms for Calderón-Zygmund Singular Integral operators*, Applied and computational harmonic analysis, 3, 120-126(1996).